Cooperative Localization for Networks with Dynamic Connectivity

Master Thesis

Konstantinos Voulgaris

Aalborg University
Electronics and IT

Electronics and IT
Aalborg University
http://www.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Cooperative Localization for Networks with Dynamic Connectivity

Theme:
Signal Processing

Project Period:
Fall Semester 2015

Project Group:
Group 971

Participant(s):
Author
Konstantinos Voulgaris

Supervisor(s):
Supervisor
Troels Pedersen

Copies: 1

Page Numbers: 100

Date of Completion:
June 1, 2016

Abstract:

This master thesis addresses the Connectivity issue of a Cooperative Localization Network. Addressing connectivity as a spatial relationship the focus is on the time evolution of this relationship.
A time independent probabilistic model is extended in the time domain by addressing the corresponding probabilistic output as a stationary distribution. The dynamic network is represented by a graph. A metric that address the time evolution of the graph is introduced. The thesis also develops a estimation strategy upon the calibration of the model can be based.
The comparison in between the proposed approach and the existing method, with respect to the proposed metric, indicated that the proposed model can capture the characteristics of the already existing model. The evaluation of the estimation strategy shown that the particular approach is suitable to fulfil the purpose of model calibration.

Preface

Th thesis is submitted to Aalborg University Denmark in fulfilment of the requirements for the Master of Science in Engineering. The thesis presents the findings of the conducted research on Connectivity of Dynamic Networks for Cooperative Localization purposes.

The work has been carried out during the time period from October 2015 to June 2016. I would like to express my gratitude to Professor Troels Pedersen for the structured guidance and motivative support throughout this work.

Aalborg University, June 1, 2016

Author
<kvoulg14@student.aau.dk>

Contents

Chapter 1

Introduction

1.1 Cooperative Localization

Location awareness is a feature which is essential for the development of applications in several aspects of commercial, public service and military sectors. Information collected by a mobile node though is useful only in case where the location of the node is known. A satellite navigation system termed as global navigation system is utilised by techniques, such as the Global Positioning System (GPS), in order to fulfil outdoor purposes. There are several cases though, the so–called GPS–denied environments [14], where the performance of the corresponding applications decays significantly. A very common environment where this situation arises are dense urban areas. In such cases the surroundings of the nodes i.e tall buildings affect the quality of the information exchange between the different agents. As a result numerous of applications are unable to fulfil their purposes.

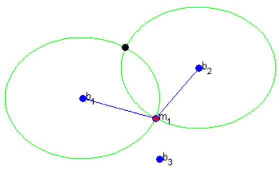
An example of the importance for high quality information exchange can be considered a rescue crew in case of an earthquake. Assuming that the members of the crew are in the area of interest but due to the presence of the fallen buildings the available equipment fails to provide reliable information regarding the location of the crew. As a result the rescue mission can be hardly coordinated and the task of guiding the crew to the target becomes very difficult. Another case related to a case of emergency can be considered the case of an ambulance in urban area where the surroundings consists of tall buildings, tunnels or narrow streets i.e in Paris. The mission of the ambulance is to recover the location of the patient as fast as possible, but due to the surroundings the task of finding the corresponding location is very difficult and the life of the patient gets in danger. Finally, a scenario which is a very common situation on a high way is the case where many tracks are stacked the one behind the other making a convoy. As a result the convoy of trucks acts like a wall to the surrounding vehicles which blocks the communication and affects the accuracy of the corresponding positioning applications. These characteristic examples indicate the there need for location–aware technologies which will overcome this kind of complications. One approach to fulfil such task it the case where each wireless node has the capability to localize by itself.

A new paradigm which aims to contribute to the direction of self localization is cooperative localization. The main scenario in cooperative localization is that each mobile device (or else node) can obtain knowledge regarding its location by exploiting information from the surrounding nodes either these are mobile or static(anchor). In order to fulfil such task, each sensor needs a processing unit and a corresponding algorithm for self-localization.

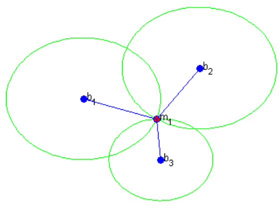
A wide range of devices trying to localize themselves, with different level of mobility such as laptops or mobile phones, could be considered. Since there is not any particular application to fulfil such task all the devices share the same goal of obtaining information regarding their location. All the devices which could provide valid knowledge regarding their location could be taken into account by the surrounding devices for the needs of cooperative localization. A main issue that should be considered in such case though, is the fact that several devices are characterized by privacy policies thus no information can be obtained regarding their position [5].

In order to represent the relation between different nodes a measure that express the distance can be used. This relation can be expressed by a range of different positioning systems. There are several techniques which are are widely used in the field of location awareness. Popular techniques applied for positioning purposes are Time of Arrival (ToA) and Received Signal Strength (RSS). Wireless networks that support radio technologies such as the WideBand(WB) and Ultra–Wideband (UWB) are well known for utilizing ToA measurements. There exist other positioning position information techniques, such as Time Difference of Arrival (TDOA), applied for localization purposes but they not considered in the current work.

A common aspect of TOA and RSS is the fact that in both cases the target node is localized by performing triangulation via the collaboration of multiple base nodes. In order to obtain an unambiguous estimation for the position of the mobile node a minimum number of three different distance measurements is required. For the case of TOA , each distance measurement is obtained by taking into account the travel time of the radio signal from a single transmitter to a mobile single receiver. On the other hand, in cases where the corresponding make use of the RSS technique the distance measurements are obtained by measuring variation on the strength of the signal from the transmitter node to the receiver node. An example that corresponds to the localization process followed by TOA and RSS is provided in figure 1.1.



(a) Potential target location if the mobile agent m_1 is one of the the two points of the intersection of the circles surrounding the base nodes b_1,b_2.



(b) Final location is the point which lies on the circle surrounding the base node b_3.

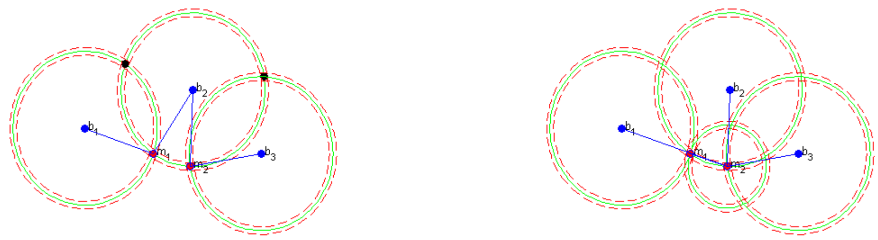
Figure 1.1: A trivial example which illustrates the localization process followed by TOA and RSS. In cases where the available measurements do not fulfil the requirement of the minimum observations the corresponding outcome is characterised by an ambiguity regarding the location of the target node like in the first subfigure. In cases where the basic requirement is fulfilled, like in the second subfigure, the applied techniques yield a single outcome.

Even though the two techniques rely on distance, their use gets complicated because of some of their characteristics. The implementation of ToA even though is a reliable technique for locating a target node, it is characterized by several drawbacks. First of all, a basic parameter which should be fulfilled in order to obtain the desired information has to do with the fact that all the nodes must be synchronized precisely [15, pp.5]. In any other case the presence of a timing error can lead to a false distance measurement. Another parameter which should be taken into account is the fact that the positions of the base nodes should be known. In any other case this knowledge is obtained through other static nodes or via GPS–equipped dynamic nodes [15, pp.5]. The implementation of the RSS technique on the other hand requires that there is previous knowledge regarding the transmitted power at some reference distance. Another drawback which is common in both cases is that for the corresponding approaches the measurements are assumed to be ideal. That kind of scenario though does not hold in the real world. Contrariwise, it is very common the case where the corresponding measurements are highly affected by the surroundings.

1.1.1 Cooperative Localization versus a non-cooperative scenario

The so far conducted analysis focus on the main reasons why the implementation of the Cooperative Localization paradigm can be beneficial for the purposes of outdoors positioning. In order to provide a more representative explanation regarding the advantages of the paradigm as well as how the methodology operates a simple example is provided in figure 1.2.

According to the localization set–up introduced in the example, three different base nodes (a_1,a_2,a_3) are attempting to localize the two mobile nodes (m_1,m_2) of the scenario. In both cases though, the requirement of the three different distance measurements is not fulfilled. In each case only two measurements are available thus the mobile target cannot be localized through triangulation. In particular the base node a_2 is unable to communicate with the mobile node m_2 while the same case holds between the base node a_3 and the mobile node m_1. This difficulty arises from the fact that the two mobile nodes are outside of the communication range of the corresponding base nodes. As a result the corresponding algorithm yields an ambiguous estimation about the position of each node since two possible locations are obtained.



(a) A non–cooperative scenario that yields an ambiguous estimation regarding the location of m_1.

(b) A scenario where the two mobile agents m_1, m_2 cooperate and an ambiguous estimation regarding the location of m_1 is obtained.

Figure 1.2: A trivial example which illustrates the benefits of the implementation of the cooperative localization paradigm. For the current scenario the main task it to estimate the location of m_1. The position of m_2 is also unknown but some position information is available.

In order to overcome this difficulty we can consider the scenario of cooperative localization. While position of the mobile devices is considered to be unknown but they can provide information regarding their position. Thus, a third distance measurement can be obtained,in such case triangulation can be applied, or any other position information that can be helpful in order to distinguish the two intersections.

1.2 Problem Analysis

So far we have introduced the basic idea that relies upon the Cooperative Localization paradigm as well as the main advantages of the approach compared to the current approaches such as the GPS. The current section aims to introduce the basic approach to form the fundamentals of Cooperative Localization to a mathematical pattern. In order to develop that kind of methodology there is the need to analyse the different aspects when considering a network.For the current work a network is generated once a connection is established in between different agents, abbreviated as nodes, participating it the same graph.Two main categories of nodes, anchors and mobiles, participate in a graph:

An anchor node is characterized by a static behaviour, meaning that the corresponding position remains unaltered over time. Therefore the coordinates of the anchor node are considered fixed for the problem.A mobile node is characterized by a dynamic behaviour which in practise means that the corresponding location varies over time. Thus there is the need to develop a framework in order to capture the dynamic behaviour of the mobile agent. The variation on the positions depends on aspects such as velocity, acceleration etc, thus the development of a mobility model should take these aspects into account.

One of the main questions that need to be answered when developing a cooperative localization framework is whether a communication link can be established or not. In order to obtain the measurements that are required so that we can track an agent a basic requirement that needs to be fulfilled has to do with the fact that the agents can exchange information. Therefore a framework is required in order to make the distinction upon the agents that are capable of having a connection and those that are not. Signal propagation is affected by the transmission environment, thus parameters such as fading or the transmission power have a severe impact on the connectivity issue. Another parameter that should be taken into account when considering the connectivity issue has to do with the dynamic behaviour of the mobile agent. In such case the distance in between the different agents participating in the network varies and as a result the impact from the transmission environment differs over time. Thus there is a need for a framework that addresses all these aspects of the connectivity problem.

Once a connection is established in between the agents the process of exchanging information begins. For each of the established links a distance observation is obtained for the nodes. The process of information exchange is characterised by several uncertainties, introduced by the signal propagation environment, thus each of the measurements may have an error. Consequently the collected data is characterised by a level of inaccuracy. Thus the mathematical pattern which is developed in order to represent the collected data should take these aspects into account.

Derived on this methodology the general mathematical pattern upon Cooperative Localization is formulated, is as follows:

- The mobility model describing the dynamic behaviour of the mobile agents.

- The connectivity model which describes the dynamic relationship of information exchange in between the agents.

- The range model describing the noise introduced in the range measurements.

Each of the different steps can be viewed as an interconnection block. In such case the relation in between the different blocks can be represented as follows:

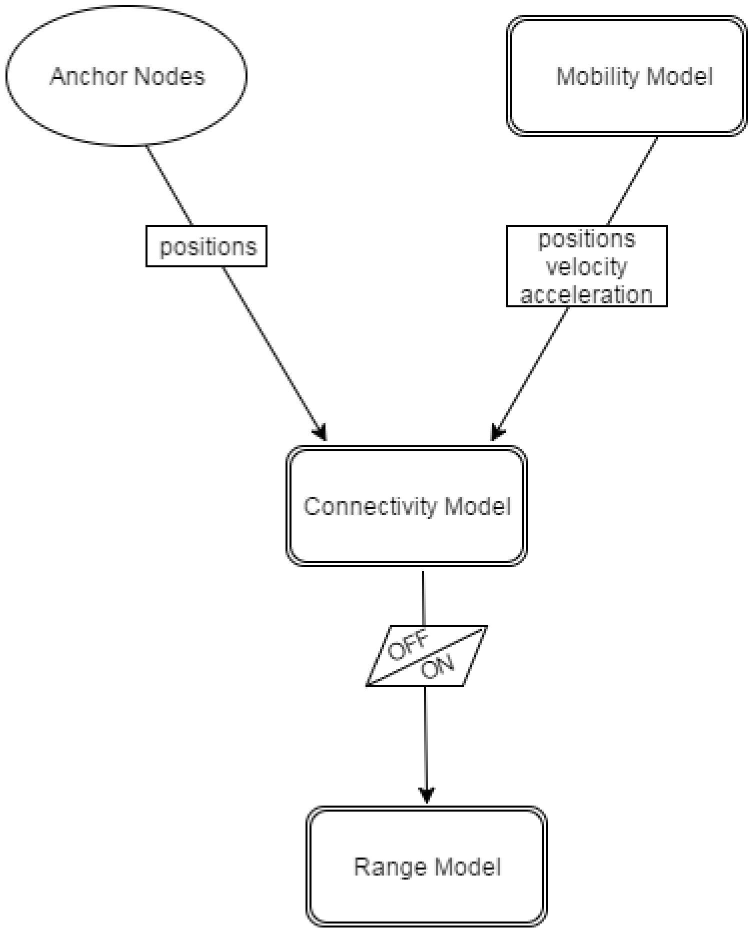


Figure 1.3: Schematic representation of the Cooperative Localization process.

The basic task of the Cooperative Localization paradigm is to provide an output regarding the location of the agent. The so far conducted analysis on the problem describes the different aspects that should be taken into account considering the nature of the problem. Once the different stages are distinguished and the corresponding mathematical patterns are developed the next step is to develop a technique that yields an output based on the these patterns. In such case the localization problem can be addressed as an estimation problem. An ordinary localization system, i.e a wireless positioning system (WPLS), provides distance measurements which in such case those could be the TOA measurements. Assuming that the model that describes the dynamic behaviour of the target is known, an estimation technique can be implemented in order to obtain an estimate regarding the location of the target.

1.3 Existing work on the field of Cooperative Localization

The research in the field of Cooperative Localization as well as the a network research in general requires a simulation pattern that provides realistic conditions so that the obtained results may not be misleading. Thus it is desired that the construction of the corresponding models is done in way so that it is close to real time circumstances.

One of the main concepts when considering a dynamic network is related with the mobility pattern that characterise the behaviour of the mobile participants. The so far conducted research in the field of Cooperative Localization that refers upon the mobility aspect provides a detailed overview regarding the different parameters such as velocity, acceleration etc. There exist several approaches upon this issue in the corresponding literature that address the problem from various points of view. Thus it is up to the designer of the application to select the one that fits better to a particular application. The interested reader could even find several studies that make a classification upon the different approaches in [**ref:mobile_research**] or in [**ref:mobile_research_1**]. The so far developed approaches though are considered to provide a fairly simple and sufficient approach to capture the basic characteristics of the current issue thus no further investigation is conducted.

Another issue that needs to be addressed as well is the one related with the inaccuracy introduced in the distance measurements. The general pattern that is followed for the current aspect of the estimation problem is similar like the one introduced in [15, pp.29]. The current approach in the existing literature has as well covered the main aspects of the problem. An obvious comment that can be discussed upon the general pattern is that it is not related with one of the main aspects of the current issue which in this case is distance. New approaches on the field of location awareness though attempt to address that issue like the one proposed in [4]. Therefore no further investigation takes place on this issue in the current thesis.

On the other hand the so far developed methods in the field of cooperative localization lack of expressing the issue of whether two agents can exchange information or not. Even though several approaches have been proposed, none of them succeeds to express all the aspects considering the connectivity aspect of the paradigm. Each of the different approaches satisfy partially the main aspects of the problem, none of them though attempts to combine all the parameters such as distance,mobility or time dependency at once. Therefore the main investigation of the current thesis focus on conducting an analysis upon the different aspects of the connectivity issue.

1.4 State of the art methods

The main task of the current thesis is to investigate the spatial relationships developed in a Network with Dynamic Connectivity. In the current work,the term network is used to describe the relationship developed in between different agents (abbreviated as nodes) aiming to exchange information with each other. We seek to develop a method describing the way that this relationship varies over time. So far there exist two proposals in the literature:

1. A Connectivity model introduced by *Savic* and *Zazo*.[12]
2. A Connectivity model introduced by *Henk Wymmersch*.[14]

A common aspect of the two proposals is the fact that both of them take into account the main parameter of the problem which in our case is distance. Beyond that though different approaches are followed to address the problem. Each of the following approaches addresses different aspects of the problem, such as transmission power or the mobility of the nodes, but in the same time they are characterized by several drawbacks.

- The main disadvantage of the proposal of *Savic* and *Zazo* is that the connectivity of the network is considered as a spatial relationship in between the different participants which is independent on time.
- On the other hand the deterministic function of *Henk Wymeersch* is a proposal that considers ideal signal propagation scenario, (ignoring e.g signal fading) which is not the case for real applications.

1.5 Problem Statement

The current work aims to overcome the impracticalities of the so far developed approaches. Therefore a connectivity model is introduced where the following parameters are taken into account:

- The dynamic behaviour of the nodes(e.g the move of mobile nodes.)
- Distance and the way that this aspect affects the connectivity in between the agents.

One of the issues considering a network with multiple communication links,which can be represented by a graph, is to evaluate the way that this relationship evolves over time. So far no such measure is proposed that can address the time evaluation over a graph. Therefore the current work also addresses the following issues:

- The time dependency of the connectivity.
- A metric that represents the time dependency of a dynamic network.

One of the main aspects when considering a model is what kind of strategy can be followed to calibrate the model. Hence the thesis another task of the thesis is the following:

- Develop a calibration strategy for the proposed model.

1.6 Thesis outline

The thesis is organized as follows:

Chapter 2 The general notation of the network is presented. Chapter 2 also introduces mobility model that describes the dynamic behaviour of the nodes as well as the model upon the distance measurements are obtained.

Chapter 3 A presentation of the state of the art methods that address the connectivity issue. An analysis upon the advantages and disadvantages of the two methods is conducted alongside with the main aspects of the problem.

Chapter 4 A *Markovian* model to address the connectivity issue is proposed and a topological analysis along with an analysis on the time domain are conducted.

Chapter 5 The *Markovian* model is dual paramteric. This chapter proposes methods upon model calibration can be based..

Chapter 6 A metric that evaluates the correlation of a dynamic network over time is introduced. A review of the performance of the model as well as upon the performance of the methods for obtaining the parameters of the model takes place with respect to the proposed metric.

Chapter 7 This chapter makes the conclusion of the thesis and presents the outcomes from the conducted research.

Chapter 2

Notation Topology, Mobility model and Range Error model

The current chapter introduces a number of sets which describe the topology of the network. The notation will be used throughout the master thesis. We also introduce the methods upon which address issues such as the mobility of the nodes and the model upon the noisy measurements are obtained.

2.1 Network topology and notation

For each wireless device that participate in the topology of the network a vertex, abbreviated node, is added in a graph and for each of the established communication links an edge is drawn. To illustrate this the simplest cooperative localization scenario is used as an example and it is presented in figure 2.1.

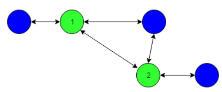


Figure 2.1: An illustration of a simple cooperative localization scenario with a topology consisting of three nodes. Three anchor nodes marked with blue and two mobile marked with green and specified by a unique number.

The given example illustrates a particular time instance n of the network topology. The network topology varies over time thus there is a need for an index to specify each time instance of the corresponding graph. For $n = 0$ the network topology is generate while N is a finite time elapsed. Hence , we state:

$$n \in N \equiv \text{discrete time index of the dynamical system} \tag{2.1}$$

The nodes that participate in a graph are characterised by different properties. Specific categories of nodes participate in a graph and they are labelled as follows:

- $m \equiv$ mobile node
- $a \equiv$ anchor node

Based on the current categorization the different categories of nodes are grouped in disjoint sets as follows:

- $U_m(n)$ set of mobile nodes in network at time n.
- $U_a(n)$ set of anchor nodes in network at time n.

At any given time instance n the number of nodes participating in a graph are represented as follows: $U(n) = U_m(n) \cup U_a(n)$.

Each edge drawn in the graph represents a communication link established in between the agents. The set of edges expressing all the links established in a graph in a particular time instance is expressed as follows:

$$E(n) = \{(r,t) : a_{e_{(r,t)}}(n) = 1\}. \tag{2.2}$$

where $a_{e_{(r,t)}}(n)$ is an link indicator defined as follows:

$$a_{e_{(r,t)}} = \begin{cases} 1 & \text{connection} \\ 0 & \text{no connection} \end{cases} \tag{2.3}$$

Once a communication link is established in between the agents they start to exchange information. In case of cooperative localization the information take the form of noisy distance observations. The set containing all the available distance observation at time slot n si defined as:

$$D(n) = \{(r,t) : (r,t) \in E(n)\}. \tag{2.4}$$

2.2 Dynamic behaviour of the mobile node

A key aspect in a network research is the simulation pattern upon this survey is based. The characteristics of the network as well as the measuring performance are severely affected by the simulation environment. Therefore, it is important to develop or follow simulation patterns which are as close as possible to real conditions. One of the main parameters, considering a dynamic network, is the behaviour of the mobile agents. There exist several models to fulfil such purpose in the corresponding literature. Several studies have taken place on the field in order to classify the developed models. Such a survey can be found in [2] , [8] and [6].

The different studies classify the performance of the mobility models by taking into account several criteria. Thus it is crucial to specify the criteria upon the selection of the corresponding model is made. The mobility model is designed to describe the movement pattern of the mobile agents, and the way that the corresponding location , velocity and acceleration vary over time. [6, pp.2]. Complicated models exist which might focus on a specific aspect of the dynamic behaviour. Therefore there is the need for a model that can capture the deviation in the general mobile pattern followed by an agent.

A dynamic system consists of a various number of mobile and static agents which form an instant network. The mobile agents travel all around the area of interest freely and they leave or enter the area by following an arbitrary fashion. As a result the network topology varies rapidly over time. Additionally the speed of the object is affected by different factors such as acceleration, obstacle avoidance, etc. Considering these aspects, there is a need for a mobility model that can capture these characteristics so that the overall evolution of the network topology is as representative as possible [3].

The so far conducted research on the field [2], has shown that the mobility models which are characterised by a memoryless behaviour (i.e the next state has no dependency with the states that preceded it) , create unrealistic patterns. Therefore another criterion that needs to be fulfilled is that the next state of the object should have a level a dependency in between the future state of the object and the states that preceded it. Considering the nature of the problem, a fair assumption that can be made is that the next location of the mobile agent should depend on the current state of the agent or in other words the corresponding mobility model should meet the first order *Markov* property. The particular property implies that $x(n-2) \perp\!\!\!\perp x(n)|x(n-1)$. In order to provide a better understanding of this relationship, a graphical representation is introduced in figure 2.2.

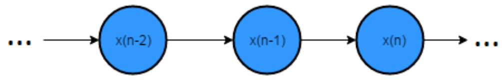


Figure 2.2: An illustration of process that meets the *Markov* property. The current state of the process depends only on the state that preceded it and is independent with the past.

In the current work did not consider any novel mobility model, but on the criteria mentioned above, we adopt the mobility model introduced in [1]. The reason why we make this choice is due to the fact that is fulfils the *Markov* property. Another advantage of the model it is structured on a simple manner thus it can be easily adopted by the connectivity model.

2.2.1 General form of the mobility model

In the following we make a presentation of the mobility model [1] that describes the dynamic behaviour of the mobile node over the graph. The new state $x(n)$ of the agent is expressed with respect to the state $x(n-1)$ that preceded it along with some process vector $z(n)$. The process vector is used to capture the variation of the position depending on parameters such as velocity or acceleration. Based on this characteristics the general form of the mobility model can then be defined as:

$$x(n) = F \cdot x(n-1) + z(n). \tag{2.5}$$

The F matrix is the transition matrix whereas $x(n-1)$ is the preceding state vector. The form of the transition matrix depends on the order of the mobility model and in particular the parameters (such as position, velocity) included in the state vector. The process vector $z(n)$ can be furthermore expressed as follows:

$$z(n) = G \cdot q(n). \tag{2.6}$$

where $q(n)$ is modelled as:

$$q(n) \sim (\underline{0}, \sigma_z^2 \cdot I). \tag{2.7}$$

The process noise $z(n)$ is a mapping of the WGN vector $q(n)$,with the corresponding covariance matrix, which is transformed by the matrix G.

Derived on this, the process noise then is distributed as:

$$z(n) \sim (\underline{0}, \Sigma), \tag{2.8}$$

where the process covariance matrix can be expressed as:

$$\Sigma = \mathrm{E}[G \cdot q(n) \cdot q(n)^T \cdot G)] = \sigma_z^2 \cdot G \cdot G^T \tag{2.9}$$

First order mobility model

For the first order mobility model, the state vector $x(n)$ represents only the position of the node for the current time instance. The variation on the position of the agent over time relies only on the variation of the speed.Thus $x(n)$ can be expressed as follows:

$$x(n) = p(n) \in \mathbb{R}^2, \tag{2.10}$$

where $p(n)$ represents the coordinates of the node on the 2–D grid.

The parameters that contribute at the state equation can be obtained by taking into account the equation introduced in [1]. Derived on this, the corresponding matrices take the following form:

$$F = I, G = \Delta t \cdot I, \tag{2.11}$$

Hence, the process covariance matrix can then be expressed as follows:

$$\Sigma = \sigma_z^2 (\Delta t)^2 \cdot I. \tag{2.12}$$

.

Second order mobility model

The so far conducted research on the field has shown that by including velocity in the states, a more realistic node movement trace is created ([2, pp.7]). A basic advantage of the current model is the fact that can be easily be extended to a second order form where the variation on the location relies on the acceleration of the mobile agent. The state vector $x(n)$ is expressed as follows:

$$x(n) = \begin{bmatrix} p(n) \\ v(n) \end{bmatrix}, \tag{2.13}$$

where $v(n) \in \mathbb{R}^2$ represents the velocity.

The components of the corresponding linear equation can be obtained by discretizing the first order differential equation introduced in [1, pp.92]. Derived on this, the corresponding matrices can then be expressed as:

$$F = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \otimes I, G = \begin{bmatrix} \frac{\Delta t^2}{2} \\ \Delta t \end{bmatrix}, \tag{2.14}$$

where $\otimes$ denotes the Kronecker product.

The covariance matrix of the process vector can then be expressed as:

$$\Sigma = \sigma_q^2 \cdot \begin{bmatrix} \frac{\Delta t^4}{4} & \frac{\Delta t^3}{2} \\ \frac{\Delta t^3}{2} & \Delta t^2 \end{bmatrix} \otimes I \tag{2.15}$$

2.3 Range error model

The estimate of the location of a mobile object relies on the knowledge of the dynamic behaviour of the object as well as the available range measurements. Each of the available observations though is characterised by a level og inaccuracy. Several factors may occur and affect the inaccuracy of the measurements which reflect the impact from the propagation environment.

In cases of dense urban areas the surroundings of the object have a severe impact on the corresponding accuracy. A possible reflection on a high building , an obstacle a neighbour of the agent etc, corrupts the information obtained by the corresponding signal. Even in cases where the agent travels on a plain terrain free of obstacles, the available information is also inaccurate. Other reasons such as ionospheric scintillation, hardware delay, thermal noise etc [4] also have an impact to the obtained measurement.

Derived on these, there is a need for a mathematical model for the error introduced in the measurements. The current work does not consider a novel model upon this issue but instead we follow the model introduced in [15, pp.29-30] where the measurement is represented as follows:

$$d_{rt} = ||p_r - p_t||_2 + w. \tag{2.16}$$

where e corresponds to the noise introduced in the measurement. For each of the measurements the corresponding error is considered to be independent to the rest of the available measurements. By taking into account these considerations e is modelled as follows:

$$w \sim N(0, \sigma_w^2) \tag{2.17}$$

2.4 Simulation scenarios

The current work investigates the connectivity status and how that varies over time. The simulations are conducted by taking into account different simulation environments depending on the number of the agents participating in the corresponding graph. In many cases trivial scenarios are investigated where the connectivity status of a pair of agents is considered. For the cases where a group of nodes, anchor or static, participate in the graph a specific simulation pattern is followed.

The basic simulation concepts are the following:

Static two agents connectivity scenario: The current simulation scenario investigates the connectivity status in between a pair of agents while they remain static over time. The main reason why we select the particular set up has to do with the fact that we want to see the the influence one of the basic aspects of the problem, in such case time, affects the exchange of information.

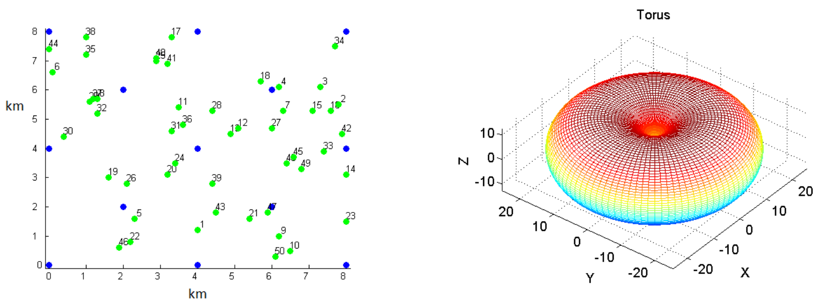
Dynamic two agents connectivity scenario: At the current set up we consider a pair of agents where at least one of the nodes has a dynamic behaviour over time. The mobile agent travels all around the simulation area without considering any boundary. The current set up provides provides information regarding the sensitivity of the connectivity status when except time, the variation of another aspect of connectivity, in such case distance, affect the connectivity in between the agents.

The trivial extended network connectivity scenario:The simulation environment spans midrange 8 km $\times$ 8 km urban area. A number of 13 anchor nodes are placed all over the simulation area in a way such that it is easier to form triangulation. On the other hand the number of agents that travel all over the simulation area is up to 50. The environment takes into account different cases for the mobile agents thus the speed variance σ_z^2 differs from one object to the other.

It is desired for the nodes to remain in the simulation area without violating the borders. Due to the dynamic behaviour of the nodes over time such case in almost unavoidable. Thus there is a need for a method to avoid such case. One approach that constrains the positions of the agents inside the simulation area is the one that introduces walls on the boundaries of the area. In such case though the trajectories of the nodes are affected which reflects to an interference to the mobility model. Therefore there is a need for a method that allows to the mobility model to describe the trajectories of the nodes without any external intervention.

Based on this desire, we choose in our study to transform the simulation area into a torus. In such case the agent continues travelling and reappears on the other side of the simulation area as soon as it reaches one boundary of the simulation area. In order to obtain a visual inspection about the form of the floor plan and how the position evolves with respect to the toroid shape an illustration over two consecutive time instances is provided in figure 2.3.

Cooperative localization paradimg aims to overcome the impracticalities introduced by the so far applied technologies which mainly arise in dense urban areas. The current set up provides information regarding the way that the connection in between multiple agents varies over time in an environment relevant to a possible application of the paradigm.



(a) Simulation region in 2–D

(b) Transformation of the simulation region into a torus

Figure 2.3: An illustration floor plan where the simulation scenario takes place. The blue points correspond to the location of the anchor nodes where the green nodes illustrate the mobile nodes.Each of the agents has is labelled with a number used as an identifier.

Chapter 3

State of the art methods

One of the main aspects in the field of location awareness is the one related with the connectivity status in between the agents participating in a dynamic network. The particular aspect is the one that indicates whether a communication link can be established in between the agents or not. The following chapter presents the two basic approaches utilized to address the particular issue for Cooperative localization purposes. The first approach, introduced by *Savic* and *Zazo*, proposes a probabilistic model to address the problem whereas the method proposed by *Henk Wymeersch* uses a decision rule to define whether the nodes can exchange information or not. One of the main factors that affect the connection between the nodes is related with the distance between the nodes. The common aspect in the two methods is the fact that distance is considered as a basic parameter to address the connectivity issue denoted as R. For the current work the distance parameter of the probabilistic model is denoted as R whereas for the rule of *Henk Wymmersch* is expressed as R_H to avoid any possible conflict. In case of the connectivity model proposed by *Henk Wymeersch* distance is the only parameter upon the decision is made whereas in case of the model of *Savic* and *Zazo* exploits the probability of having a connection over distance. The current chapter aims to provide a deeper analysis regarding the characteristics of the models and highlight the advantages and disadvantages of each method.

3.1 Connectivity model by Savic and Zazo

The model proposed by *Savic* and *Zazo* introduces a probabilistic approach to address the problem of the spatial connection between two nodes. The probabilistic model takes the form of a function similar to the one that corresponds to the probability density function (pdf) of a *Gaussian* distribution. Unlike the *Gaussian* pdf though, the corresponding input comes only from the non–negative side of the x axis thus the graphical representation of the function in the 2–D grid looks like a representation of a one sided *Gaussian* pdf. The probabilistic function introduced by *Savic* and *Zazo* expresses the probability of establishing a connection in between the agents as a ratio in between the distance that interpolates in between the agents and a range parameter R that controls the probability. Hence the probabilistic function reads:

$$P(E(r,t) \in U) =: \exp(-||x_t - x_r||^2/2R^2). \tag{3.1}$$

One of the main advantages of the probabilistic approach is that states the idea that connection in between the agents is characterised by a level of ambiguity besides the interpolated distance. The Cooperative Localization paradimg arises from the necessity for a method that overcomes the impracticalities introduced by the signal propagation environment to the so far developed technologies i.e GPS. In a real time environment signal propagation is affected by several parameters, such as transmission power fading etc. Hence addressing the connectivity issue from a probabilistic approach introduces a method where those parameters are also considered. The ambiguity for having a connection though is also affected by the settings of the applied technology. In cases though where for i.e the transmission power is high, the influence from the signal propagation environment is less than in cases where the opposite condition regarding the transmission power. Thus the probability for having a connection is not the same for different set–ups. Thus there is a need for a distinction in between the different settings. Therefore the control parameter R is also an advantage for the model since it can be interpreted as a parameter can express different characteristics of a possible set–up such as transmission power. The current approach is also characterized by several properties that fit very well with the nature of the problem. These properties are:

1. $\lim_{||x_t - x_r|| \to 0} P(x_t, x_r) = 1$. Meaning that while the mobile node gets closer to the transmitter the probability converges to 1. In practice the model states the idea that when the nodes are very close the one with the other it is very likely that they can exchange information since the affect from the transmission environment is low.

2. $\lim_{||x_t - x_r|| \to \infty} P(x_t, x_r) = 0$. This property indicates that when the agents are far distant from each other the transmission power is very low so the interference from the signal propagation environment severely affects the signal path and as a result the connection is more likely to fail.

3. Another property of the model is that in cases where the node returns to a position that has already been before the corresponding probability is the same. The model is only distance dependent. The current property states the idea that if for example (in case of an anchor node and a mobile node) the mobile node returns in the same position then the probability of having

a connection is the same as in the first time. However there is no guarantee of having a connection there again.

4. No time–evolution is included in the model. Considering the case where a connection is established in between the nodes, there is no guarantee that the connection is on for the next i.e nanosecond. The probability is drawn independently thus, according to the model, a different connectivity status could be established in a very short time period.

5. Assuming that the node remains static in its position the probability does not change. Considering the nature of the problem, the current property expresses the idea that when the user is in a place where he has the connection it is safe to stay on the current location so that the particular connectivity status is retained .

6. Due to the time–independence of the model another assumption that can be made is that once a connection has been established then the connection may last for an infinite amount of time.

The current model even though it proposes a more realistic approach to express the connectivity status between two nodes compared to the one proposed by *Henk Wymeersch*, it is also characterized by several obvious drawbacks. The main disadvantage of the model is that it expresses connectivity as a static relationship between the nodes without taking into account the dynamic behaviour of the nodes and also whether the nodes were related in the past or not. Another drawback is the fact that considering for example a scenario which includes a base node and a mobile node, in case where the mobile node remains static (i.e a car when the traffic light is red) the probability remains the same regardless the time instance. In such case it must be assumed that the impact of fading is constant over time thus the probability should not be altered. That kind of assumption is not fulfilled in a real time environment. In order to obtain a better idea regarding this consideration, a simple experiment is conducted. The simulation set–up considers the static agents connectivity scenario. In order to determine of either the two nodes can be connected or not the probability obtained by the model of *Savic* and *Zazo* is compared with a time independent value drawn from a uniform distribution which it is assumed that it represents fading. The results of the experiment are introduced in figure 3.1.

The experiment takes into account ten different time instances. As it can be seen from the results the connectivity status varies over the considered time period. Therefore it is unrealistic to express a relationship, such as connectivity, characterised by a dynamic behaviour over time as a static relationship in between the agents which in this case corresponds to the probabilistic output of the model.

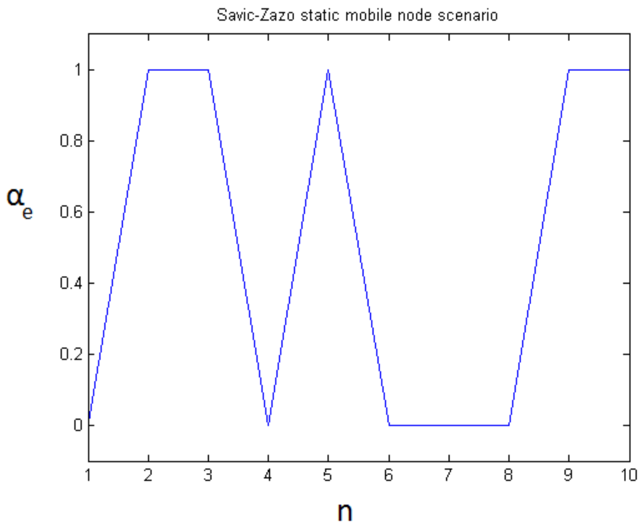


Figure 3.1: A static scenario ,considering the coordinates of the node, over different discrete time instances.a_e corresponds to the connectivity status for each instance

3.2 Connectivity model by Henk Wymeersch

The second approach introduced my *Henk Wymeersch* uses a simple rule to define whether the two agents can exchange information or not. The only parameter upon the distinction in between the different connectivity status is made is the interpolated distance in between the nodes. According to the particular approach a connection can be established in between the agents in cases where the interpolated distance does not exceed the corresponding transmitting radius R_H. The mathematical formulation of the rule is then expressed follows:

$$(r,t) = \begin{cases} \in \mathcal{E} & \text{if } ||x_t - x_r|| \leq R_H \\ \notin \mathcal{E} & \text{otherwise} \end{cases} \tag{3.2}$$

The main advantage which characterizes the particular approach is the fact that it can be easily adapted by the dynamic behaviour of the nodes. This can be done by simply inserting the new positions of the nodes with respect to the corresponding mobility model. Another property of the model is that, considering that the nodes remain static, the connectivity status remains the same. In other words, if for example we have a mobile phone which has a good connection to a particular position we do know that if we stay in the same position we will keep having a connection. Also, in case that we return in the same position we are aware of the kind of connectivity we will have. Considering this situation from a real life point of view this can be interpreted as the following example: a user is in the living room and his device is connected to the network, when the user reallocates to the kitchen the connection to th network drops. In such case the user is aware that if he returns in the location where the device was connected to the network a a communication link will be established again.

On the other hand, that kind of approach is characterised by some obvious drawbacks. The main disadvantage of the connectivity model proposed by *Henk Wymeersch* is that the affect from the signal propagation environment, such as fading, is not taken into consideration. In such case ideal conditions for signal propagation must me assumed, in the sense that the transmission power constantly overcomes the impact from the signal propagation environment. Such an approach could be followed in cases where we are considering a plane terrain without high buildings or materials which can affect the signal propagation or in general an environment that does not affect the signal path. Cooperative localization paradigm mainly focus in areas where the impact from the signal propagation environment is relatively high thus an alternative approach is required in order to overcome the impracticalities introduced to the far developed technologies. Another characteristic of the model is the behaviour of the model at the boundary of the transmission radius. For $||x_t - x_r|| = R_H$ we can be sure that a connection can be established, whereas when $||x_t - x_r|| = R_H + \epsilon$ (where ϵ is a very small value) the connectivity drops. The connectivity status is a very sensitive issue and it is quite possible that the connectivity might drop by even making a small step. The disadvantage of the particular approach though is the level of certainty in both cases. By seeing the issue from a probabilistic approach it is unrealistic for the probability to drop from 1 abruptly to zero.

3.3 Comparison between the two approaches

The so far analysis on the different approaches mainly focus on the characteristic properties of the methods and provides an intuition regarding the advantages and disadvantages of each method. The current subsection aims to provide a deeper analysis upon the behaviour of the two approaches based on simulation scenarios which represent trivial cases considering the connectivity issue. In that way we aim to highlight the differences im between the different methods in a more practical manner. Therefore a variation of the dynamic agents connectivity scenario. In that way we seek to illustrate the sensitivity of the connectivity issue even in cases when only one of the agents travels in the simulation area. The connectivity status in case of the model of *Savic* and *Zazo* is defined by making a comparison between the corresponding probability and a value, different over the corresponding time instances, which is drawn from a uniform distribution. For the needs of the simulation R and R_H are equal even though the two parameters represent different parameters. Along with the illustration connectivity status determined by each method, an illustration of the variation of distance over time instances is also provided, so we can examine the influence of this aspect to each of the methods. The obtained results are introduced in figure 3.2.

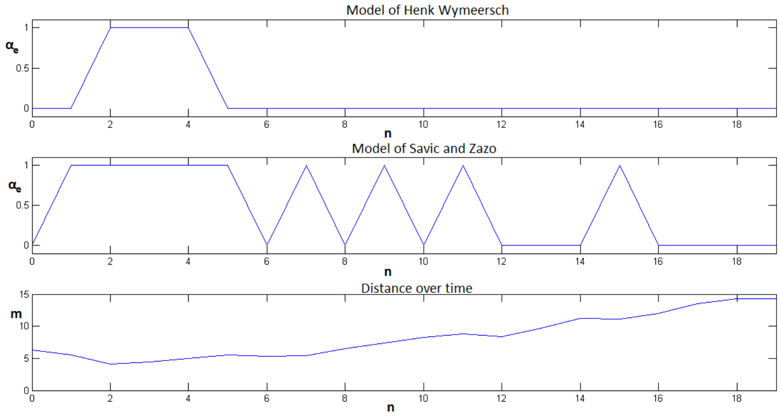


Figure 3.2: Comparison between the models. The first figure illustrates the connectivity status with respect to the model proposed by *Henk Wymeersch* , the second figure illustrates the connectivity status with respect to the model proposed by *Savic* and *Zazo* whereas the third figure illustrates the variation of the distance expressed in meters (m).

The results indicate that one of the main differences in between the different approaches is the behaviour of the methods over time. The connectivity model of *Henk Wymeersch* yields a flat output depending strictly on distance whereas the output from the probabilistic model tends to fluctuate over time. This kind of result is expected since the model of *Savic* and *Zazo* is characterised, due to the probabilistic structure, by uncertainty regardless distance whereas the connectivity model of *Henk Wymeersch* yields a constant outcome based strictly to distance. In order to provide a more representative illustration of the differences the location of the mobile node along with the corresponding connectivity status are illustrated for each model in figures 3.3 and 3.4 ccordingly.

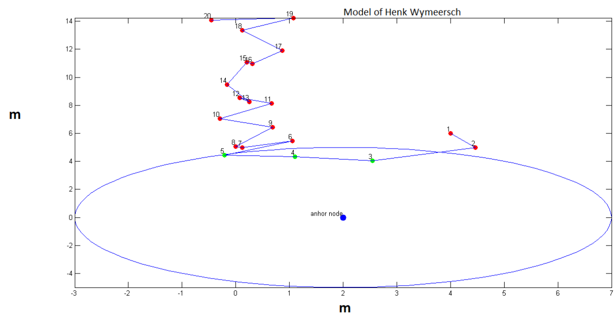


Figure 3.3: Positions and connectivity status with respect to the model of *Henk Wymeersch*.

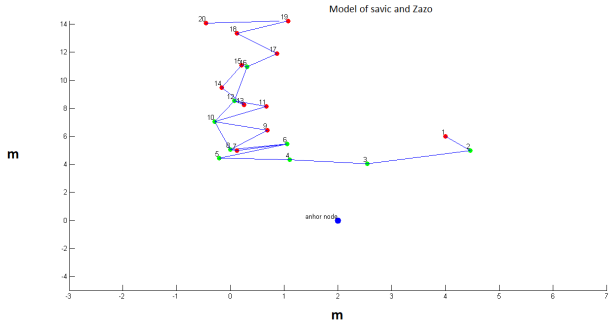


Figure 3.4: Positions and connectivity status with respect to the model of *Savic and Zazo*.

The conducted simulation provides an overview regarding the connectivity status between the nodes for a limited time period. In order to obtain an overall idea about the spatial relationship over time and investigate a more realistic scenario an extended simulation takes place where 10000 different discrete time instances are taken into account. The obtained results are illustrated in figure 3.5.

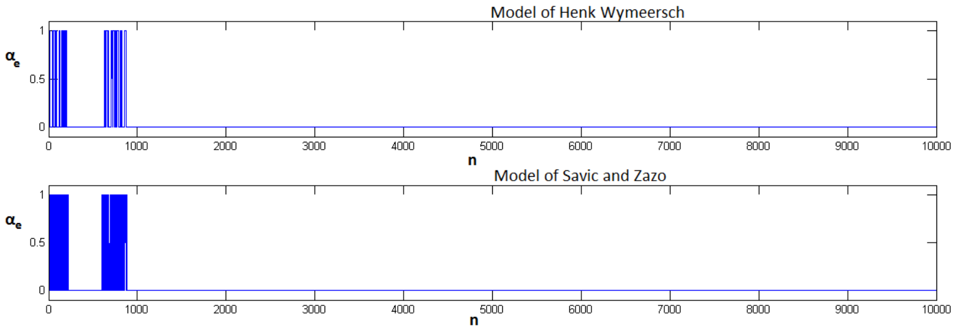


Figure 3.5: Comparison between the models. The first figure illustrates the connectivity status with respect to the model proposed by *Wymeersch* , the second figure illustrates the connectivity status with respect to the model proposed by *Savic* and *Zazo* .

The obtained results illustrate the same pattern as the one introduced in figure 3.2 but in a more extended version. the model proposed by *Savic* and *Zazo* has a more dense behaviour over time. A new aspect though is that the mobile node has the tendency to return into a region where the connection with the anchor node can be restored. The density of the results do not allow any further investigation in the current form thus the analysis will take place over shorter time periods.

Firstly the comparison in between the different methods focus on the time period where a communication link is established for the first time and ends at the time slot where the connection drops for the first time with respect to the proposal of *Henk Wymeersch*. The obtained results are illustrated in figure 3.6.

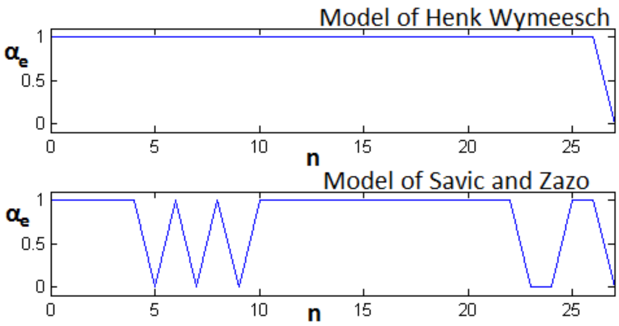


Figure 3.6: Comparison between the models. The first figure illustrates the connectivity status with respect to the model proposed by *Henk Wymeersch* , the second figure illustrates the connectivity status with respect to the model proposed by *Savic* and *Zazo*.The time period lasts until the first drop of the communication link with respect to the model of *Henk Wymeersch*.

The main task of the simulation is to provide an overview about the connectivity status over a long time period. So far we have illustrated the trivial case where the connectivity is established for the time instance until the time when the connectivity drops for the first time. At this point we will proceed by illustrating the time period when the connection is established again after the first drop. The results are illustrated in figure 3.7.

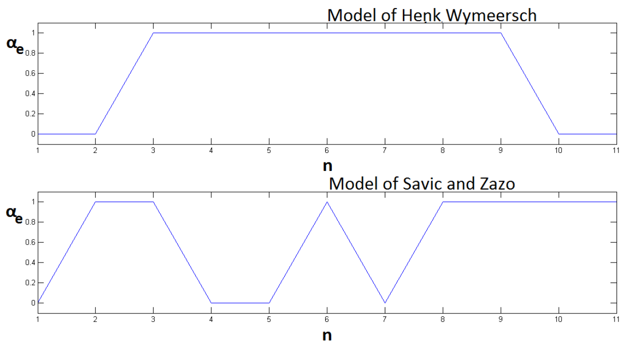


Figure 3.7: Comparison between the models. The first figure illustrates the connectivity status with respect to the model proposed by *Henk Wymeersch* , the second figure illustrates the connectivity status with respect to the model proposed by *Savic* and *Zazo*.The time period lasts from the first time the communication link is restored.

An interesting outcome from the simulation study, considering the connectivity model of *Henk Wymmersch*, is that once the connection is lost for the first time the connectivity status tends to fluctuate over time. Considering the main idea upon the the particular connectivity model is based this simply means that the agent move in the boundary of the transmission radius hence the agent exits and enters in the transmission area frequently. On the other hand the probabilistic approach is characterised by a higher fluctuation due to the fact that a longer distance inserts a higher level of uncertainty for the connectivity outcome.

The obtained results indicate that both approaches are characterised by a kind of paradox considering the nature of the problem. The agents attempt to connect in a time varying environment where nodes, i.e individuals vehicles etc, move all around changing the layout of the area upon the signal is propagated. Nevertheless the criteria in both approaches, regardless if the they are probabilistic or deterministic, remain unaltered. In case of the connectivity model of *Henk Wymeersch* the propagation environment is totally ignored whereas from the very first moment whereas in case of the probabilistic approach the affect in the propagation path in a dynamic environment is considered as static.

A new outcome from the extended simulation study is that in both approaches a state of no connection is in many cases followed by a state where a communication link is established in between the nodes. The so far followed methods though do not address this kind of issue since the structure of both the models is not state based.

3.4 Conclusion

The current chapter introduced the so far developed approaches in the field of location awareness which address the issue of connectivity. The existing methods fulfil different requirements of the problem. On one hand the connectivity model proposed by *Henk Wymeersch* is characterised by time dependency which relies on the mobility model. As long as the trajectory of the node is within the transmission radius the agents are connected with each other whereas the opposite scenario holds for the case where the corresponding trajectory is located outside the transmission area. The unrealistic aspect of the particular approach has mainly to do with the fact that the signal propagation environment is not considered as a part of the connectivity issue. On the other hand the probabilistic model introduced by *Savic* and *Zazo* proposes a more realistic approach since the affect of the propagation environment is considered as an aspect of the connectivity issue but it is also characterized by several drawbacks,i.e the dynamic change of the propagation environment and time independence. The conducted analysis indicated that a connectivity method should fulfil account some basic requirements. One of the main issues is related with the distance and the affect from the propagation environment on a time varying basis and not through a constant manner. Another issue which should be addressed by the model is the fact that there is a relationship between the current connectivity status with the one that preceded it. A very common approach to address that type of issues is to develop methods that fulfil the first order *Markov* property.

Chapter 4

The *Markovian* model

The connectivity status in between the nodes is a spatial relationship characterised by a variational behaviour over a time period. The simulation study indicated that one of the main characteristics of this behaviour is the fact that the next connectivity status is dependent to the current status regardless what happened in the past. Considering the nature of the problem this is to say that if for example a connection is established at the current time instance it is quite possible that the nodes will remain connected the following time instance regardless if no connection was established two time instances before. A very interesting result of the so far conducted analysis is also the fact that if two nodes are not connected the current time there is a strong possibility that at the next step a transaction from the status of no connection to a status of connection can occur. The common aspect of all the different scenarios though is the fact that they fulfil the first order *Markov* property. Derived on this outcome the current work proposes a *Markovian model* to address the connectivity issue. The fundamental idea about the model is based upon the proposal by *Savic* and *Zazo*, the *Markovian* approach of the mode aims to extend the original proposal in the time domain. The following chapter introduces the fundamentals upon the *Markovian* model is based starting with the theoretical analysis. Furthermore a simulation study is conducted in order to investigate the way that the original proposal adopts the updated approach. Therefore an analysis upon the behaviour of the *Markovian* model in the time domain is conducted. Finally a discussion takes place regarding the properties of the model parameters as well as a preliminary analysis about methods upon the model parameters can be obtained.

4.1 Markov chain

Therefore the connectivity status over a time period could be considered as a *Markov* chain since it meets the following requirement:

$$P(X_n = s|X_0 = x_0, \cdots, X_{n-1} = x_{n-1}) = P(X_n = s|X_{n-1} = x_{n-1}) \quad (4.1)$$

Two different cases hold for the connectivity status in between the agents, either the agents can exchange information with each other or not. Considering the two cases as different states of a *Markov* model, the first case could be indicated as S_0 whereas the second scenario could take the form of state S_1. Thus, the discrete number of states can then be expressed as a set that takes the following form:

$$S = (S_0, S_1) \quad (4.2)$$

At each time instance the connectivity status has a probability to remain at the same state the very next moment and a probability of transition to the next state. An example of this type of process is illustrated in figure 4.1

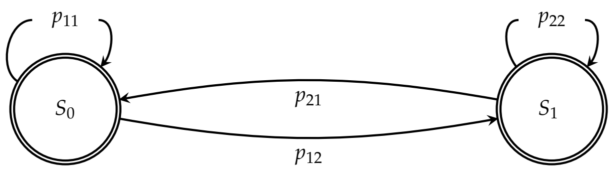


Figure 4.1: An example of a 2–state *Markov* model .

The corresponding probabilities could then be represented in the matrix form P where:

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \quad (4.3)$$

P is a stochastic matrix, which is to say that it fulfils the following properties ([7, pp.195]):

- P has non–negative entries.
- The sum over the rows is equal to 1: $\sum_j p_{ij} = 1$.

Based on the second property of the stochastic matrix, the transition matrix P can be rewritten as follows:

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}, \text{where } \alpha, \beta \geq 0 \quad (4.4)$$

Distribution over the states

The distribution over the given states can take the form of a stochastic row vector. Since the connectivity status can take only two states, the stochastic is formed with respect to the probability of establishing a communication link and the probability that the opposite case may occur. The distribution over the different time slots can then be expressed as follows:

$$\pi^{(n+1)} = [\pi_1^{(n+1)}, \pi_2^{(n+1)}] = \pi^{(n)}P = (\pi^{(n-1)}P)P = \cdots = \pi P^n. \tag{4.5}$$

Eigenvalue Decomposition

The main parameter that characterizes the distribution over the different time instances is the form of the transition matrix P. In order to obtain a better idea regarding the entries of the transition matrix and the way that the entries vary over time, a structure analysis is conducted. The first step of the analysis is to obtain the eigenvalues λ_i of P. In order to obtain the eigenvalues of P we have to solve the following equation:

$$det|P - \lambda I| = 0 \tag{4.6}$$

The solution of the equation introduced in (4.6) yields:

1. $\lambda_1 = 1$
2. $\lambda_2 = 1 - \alpha - \beta$.

Based on the obtained solution the transition matrix can be reformulated as follows:

$$P = U \begin{bmatrix} 1 & 0 \\ 0 & 1-\alpha-\beta \end{bmatrix} U^{-1}, \tag{4.7}$$

where U the matrix containing the corresponding eigenvectors.

The current form of P provides an easier way to analyse the structure of the transition matrix over the different time instances. For the $n+1$ time instance the transition matrix takes the following form:

$$P^n = U \begin{bmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{bmatrix} U^{-1} \tag{4.8}$$

The power of the transition matrix is characterised by two properties. The first property is that for $n = 0$, $P = I$ where I is the identity matrix. The interpretation of this property considering the nature of the connectivity problem is to say that if for example the two agents are connected , then if the clock does not run the agents will surely remain connected. The second property of P is that while the time elapses converges to a steady form. In order to explain why this phenomenon occurs, an analytical expression for the entries of the n–th power of P is provided.

Based on the form of P^n introduced in equation (4.8) the element which is placed in the first row and the first column can be written as follows [13] :

$$p_{11}^{(n)} = A + B(1-\alpha-\beta)^n, \text{for some A and B.} \tag{4.9}$$

For $n = 0$:

$$p_{11}^{(0)} = 1 => A + B = 1. \tag{4.10}$$

For $n = 1$:

$$p_{11}^{(1)} = 1-\alpha => 1-\alpha = A + B(1-\alpha-\beta). \tag{4.11}$$

The solution of the equations (4.8),(4.10) yields:

$$(A,B) = \frac{(\beta,\alpha)}{\beta+\alpha} \tag{4.12}$$

By following the same methodology for the rest of the components, P^n can be rewritten as follows:

$$P^n = \begin{bmatrix} \frac{\beta}{\beta+\alpha} + \frac{\alpha}{\beta+\alpha}(1-\alpha-\beta)^n & \frac{\alpha}{\beta+\alpha} - \frac{\alpha}{\beta+\alpha}(1-\alpha-\beta)^n \\ \frac{\beta}{\beta+\alpha} - \frac{\beta}{\beta+\alpha}(1-\alpha-\beta)^n & \frac{\alpha}{\beta+\alpha} + \frac{\beta}{\beta+\alpha}(1-\alpha-\beta)^n \end{bmatrix} \tag{4.13}$$

We do know though that there exists the following property:

$$1-\alpha-\beta < 1 => \lim_{n\to\infty}(1-\alpha-\beta)^n = 0. \tag{4.14}$$

Therefore the transition matrix over the different time slots converges to the following form:

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \frac{\beta}{\beta+\alpha} & \frac{\alpha}{\beta+\alpha} \\ \frac{\beta}{\beta+\alpha} & \frac{\alpha}{\beta+\alpha} \end{bmatrix} \tag{4.15}$$

This property in practise means that the transition matrix becomes independent of the time evolution.

Stationary Distribution

The row vector $\pi = [\pi_0, \pi_1]$ is called a stationary distribution of the chain if the corresponding entries fulfil the following properties ([7, pp.207]):

1. $\pi_j \geq 0$ for all j, and $\sum_j \pi_j = 1$.

2. $\pi = \pi \lim_{n\to\infty} P^n$.

Once the basic properties of the stationary distribution are stated, the next step is to investigate the form of π with respect to the form of the transition matrix when $n \to \infty$ as it is formed in equation 4.15. Therefore the second property of the stochastic vector can be rewritten as follows:

$$[\pi_0, \pi_1] = [\pi_0, \pi_1] \begin{bmatrix} \frac{\beta}{\beta+\alpha} & \frac{\alpha}{\beta+\alpha} \\ \frac{\beta}{\beta+\alpha} & \frac{\alpha}{\beta+\alpha} \end{bmatrix} \tag{4.16}$$

The system of linear equations yields:

1. $\alpha\pi_0 - \beta\pi_1 = 0$
2. $\beta\pi_1 - \alpha\pi_0 = 0.$

The obtained system of equations does not yield a unique vector as a solution since the equations are linearly dependent. In order to obtain the final form of the stochastic vector we can utilize the first property of the stochastic vector. In other words we can simply insert $\pi_1 = 1 - \pi_2$ in either of the obtained equations. The solution of the corresponding system of equations yields:

$$[\pi_0, \pi_1] = [\frac{\beta}{\beta+\alpha}, \frac{\alpha}{\beta+\alpha}] \tag{4.17}$$

Simulations

The connectivity model proposed by *Savic* and *Zazo* is characterised by having the ability to adopt the influence of several parameters when considering signal propagation environment, such as fading, signal power etc. On the other hand it is also characterised by the fact that the corresponding probabilities are independent over time. In order to overcome this impracticality a very simple approach is to combine it with two state *Markov* model. In such case the probability of having as well as the probability for not having a connection form the stationary distribution. Thus in that case the stationary vector introduced in equation (4.17) is expressed as:

$$[\pi_0, \pi_1] = [\exp(-||x_t - x_r||^2/2R^2), 1 - \exp(-||x_t - x_r||^2/2R^2)] \tag{4.18}$$

A basic component of the two state *Markov* model is the transition matrix P. The main aspect when forming P is to define α and β. Based on the form of the stationary distribution there is a single relationship in this two parameters thus there is an infinite solution of pairs which could satisfy this relationship. Therefore an arbitrary choice upon one of these parameters is made, in order to form P, and the other is selected accordingly. Based on this methodology on forming P several realizations are conducted in order to investigate the behaviour of the *Markovian* model. The obtained results are illustrated along with those corresponding to the model of *Savic* and *Zazo* in order to demonstrate the differences in between the two approaches. Therefore a simulation study takes place with respect to the static agents connectivity scenario. The current simulation set up provides an overview regarding the differences in between the the two models while the only aspect of the problem that varies is time.

There exists an infinite amount of different examples upon the comparison can be made. For the current simulation set up we consider cases of connectivity upon we can have an intuition regarding the connectivity status beforehand. Therefore we choose to illustrate the results from the following simulation set ups:

- A simulation set up where the two agents are closely spaced thus it is expected that it is more likely to have a connection.
- A simulation set up where the two agents are far distant with each other hence a connection is more likely to fail.

In both cases the obtained results from the model of *Savic* and *Zazo* are illustrated first while the results from the *Markovian* approach follow.

The time variation of the connectivity status when the mobile agent is relatively close to the base node is illustrated in figure 4.2.

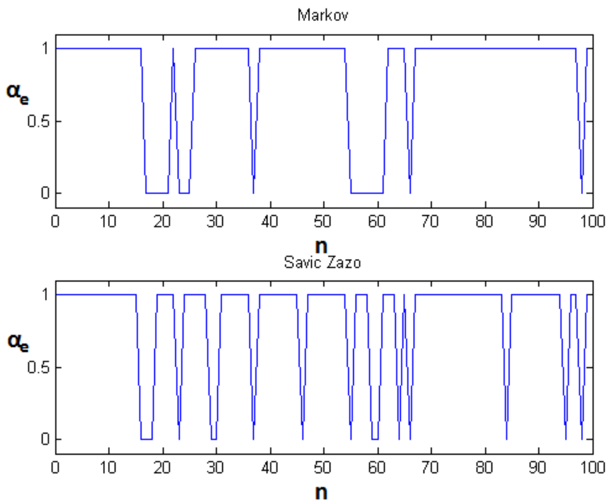


Figure 4.2: The current simulation scenario illustrates the behaviour of the model when the probability for having a connection is relatively high. The probability is the same for both the models:$\pi_0 = 0.83$. Settings for the *Markovian* model: $R = 10$,$\alpha = 0.05$,$\beta = 0.27$.

The results from the simulation scenario which investigates the variation on connectivity when the agents are relatively far the one from the other are illustrated in figure 4.3.

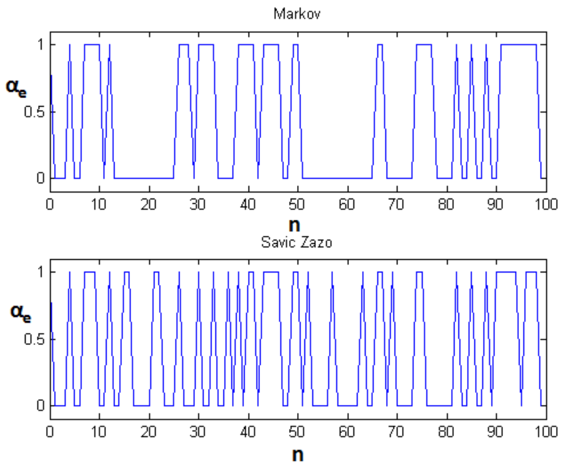


Figure 4.3: The current simulation scenario illustrates the behaviour of the model when the probability for having a connection is relatively low. In both cases the probability is the same :$\pi_0 = 0.48$. Settings for the *Markovian* model: $R = 10$,$\alpha = 0.31$,$\beta = 0.3$.

The main conclusion that can be made based on the obtained results is that the *Markovian* model yields a more stable behaviour over time. That kind of tendency is expected since the next state depends on the state the preceded it. Thus in cases where for example there exists a connection for the current state it is more possible that a similar case can repeat from the next state compared to a scenario for having a connection in the future while no connection is established in the present. The same case holds for the scenario where the current state has no connection established. Generally, this kind of pattern fits the nature of the problem. Considering the case where a connection is established in between two agents at an arbitrary time instance it is quite possible that the agents will remain connected the very next moment so they can exchange information. A general outcome that can be obtained with respect to the simulation results is that considering the connectivity as a problem we do know that the connectivity status, regardless if the connection is established or not, has in general a flat response over time. The transitions from one state to the other occur but in general the do not follow a consecutive pattern as the one introduced by the original model.

Time Correlation

The *Markovian* approach inserts time as an aspect of the connectivity status. One of the main issues that characterise time is the correlation among the different time instances. Considering the nature of the problem, the correlation among the different time instances can be interpreted as the probability of having the same connectivity status over a time period of N. The first step when conducting such an analysis is to investigate what is the average outcome for the connectivity status over a long time period. For the *Markovian* model the expectation reads:

$$\mathbb{E}[a_e] = 1 \cdot \pi_0 + 0 \cdot \pi_1 = \pi_0 = \frac{\beta}{\beta + \alpha} \tag{4.19}$$

We seek to investigate the case where a pair of agents remains connected over a long time period. That kind of relationship can be expressed as follows:

$$\begin{aligned} R_{a_e}(k) := \mathbb{E}[a_e(n+k) = 1 | a_e(n) = 1] = \mathbb{P}(a_e(n+k) = 1 a_e(n) = 1 \text{ and } a_e(n) = 1) \\ = \mathbb{P}(a_e(n) = 1)\mathbb{P}(a_e(n+k) = 1 | a_e(n) = 1) \end{aligned}$$

For $k = 1$, with respect to equation (4.17) and the form of P introduced in (4.4) the probability for having a connection the moment right after the connection is established is expressed as:

$$\begin{aligned} \mathbb{P}(a_e(n) = 1 \text{ and } a_e(n+1) = 1) = \mathbb{P}(a_e(n) = 1)\mathbb{P}(a_e(n+1) = 1 | a_e(n) = 1) \\ = \frac{\beta}{\beta + \alpha} \cdot (1 - \alpha) \end{aligned}$$

For $k = 2$, we want to obtain the probability for having a connection two moments after the connectivity is established for the first time. Due to the *Markov* property a state depends only on the state that preceded it and it is independent to the past, in our case this means that: $(n+2) \perp\!\!\!\perp n \cap (n+2) | (n+1)$. Therefore the probability is expressed as follows:

$$\begin{aligned} \mathbb{E}[a_e(n+2) = 1 | a_e(n) = 1] = P(a_e(n) = 1 \text{ and } a_e(n+2) = 1) \\ = \mathbb{P}(a_e(n) = 1)\mathbb{P}(a_e(n+2) = 1 | a_e(n+1) = 1 \text{ or } a_e(n+1) = 0) \\ = \mathbb{P}(a_e(n) = 1)(\mathbb{P}(a_e(n+2) = 1 | a_e(n+1) = 1) \cup \mathbb{P}(a_e(n+2) = 1 | a_e(n+1) = 0)) \\ = \frac{\beta}{\beta + \alpha} \cdot ((1 - \alpha)^2 + \alpha\beta) \end{aligned}$$

The same procedure can be followed for the rest of the cases in order to obtain the probability for the N different instances. A more simple approach though can be followed in order to obtain the correlation among the different time instances. This can be done by exploiting properties of the transition matrix. Based on equation (4.5), the distribution for the $n+k$–th distance can be written as:

$$\pi^{(n+k)} = \pi^{(n)} \cdot P^k. \tag{4.20}$$

The probability for having, after k time instances connection again can then be expressed as:

$$\mathbb{E}[a_e(n+k) = 1 | a_e(n) = 1] = \pi_0^{(n)} \cdot P_{11}^k. \tag{4.21}$$

Generally each of the entries of the k–th power of P expresses all the possible transitions starting from state i and ending to state j over a time period of k. So for example the P_{11}^k component expresses all the possible transitions from a state of connection to a state of connection while P_{12}^k expresses all the possible transitions from a state of connection to a state where connectivity is lost and so on.

Based on the form of P from equation (4.13) as well as the form of the stationary vector from equation (4.17) the form of correlation introduced in equation (4.21) can be rewritten as follows:

$$\mathbb{E}[a_e(n+k) = 1 | a_e(n) = 1] = \frac{\beta}{\beta+\alpha} \cdot (\frac{\beta}{\beta+\alpha} + \frac{\alpha}{\beta+\alpha}(1-\alpha-\beta)^n). \tag{4.22}$$

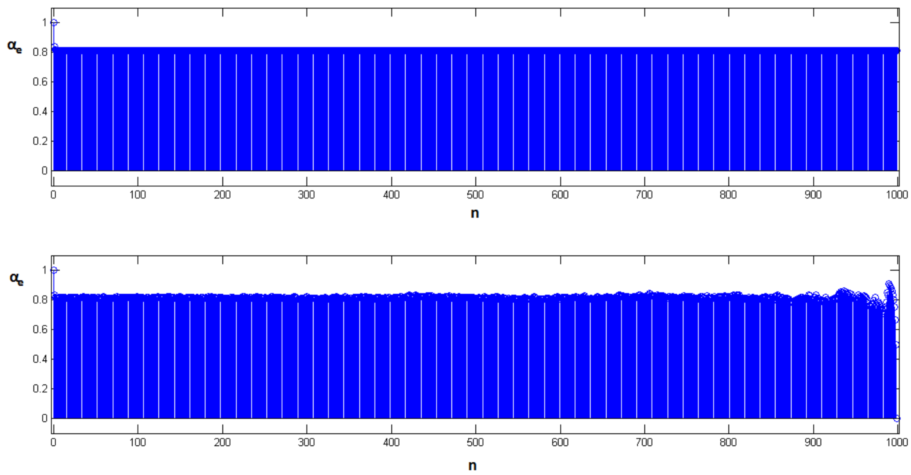
For a very large number of k the correlation converges to a steady state which is expressed as:

$$\lim_{k\to\infty} \mathbb{E}[a_e(n+k) = 1 | a_e(n) = 1] = \frac{\beta}{\beta+\alpha} \cdot (\frac{\beta}{\beta+\alpha}) = \frac{\beta^2}{(\beta+\alpha)^2} = \mathbb{E}^2[\alpha_e]. \tag{4.23}$$

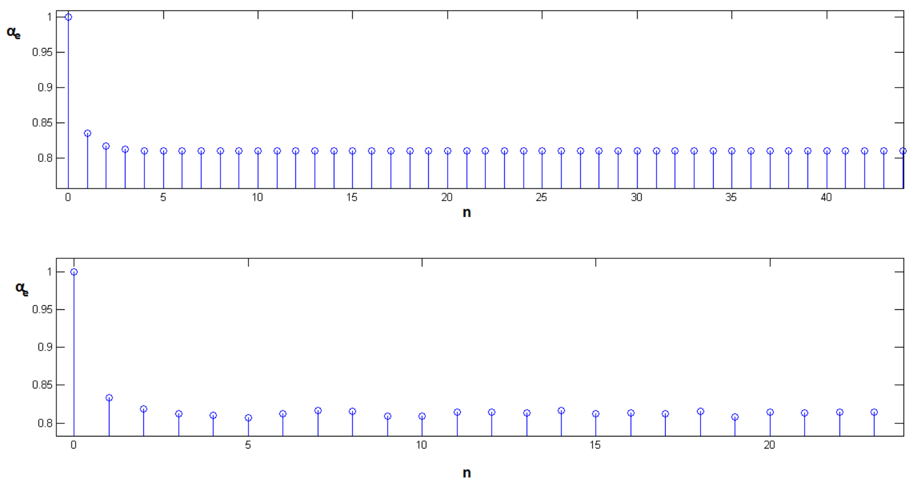
Based on these results a simulation is conducted with respect to the static agents connectivity scenario. The simulation set up considers two characteristic cases of connectivity: i) a case where the probability of having a connection is very high ii) a case where the connectivity is quite likely that it will drop. Along with the results from the theoretical derivation of the time correlation an illustration of the numerical estimation of the time correlation is also provided. The numerical estimation of the ACF is obtained by making use of the unbiased ACF estimator [11]:

$$\hat{R}_{a_e}(k) := \begin{cases} \frac{1}{N-k} \sum_{n=0}^{N-k-1} X(n)X(n+k) 0 \leq k \leq N-1 \\ \hat{R}_{a_e}(-k) - N + 1 \leq k \leq 0 \\ 0 \text{ otherwise} \end{cases} \tag{4.24}$$

The simulation takes place in a prolonged time period thus the overall results are illustrated in a very dense manner. In order to provide a closer look to the results, a zoomed in illustration of the first few instances is also provided. The results from a case where the two agents are closely spaced are illustrated in figure 4.4.



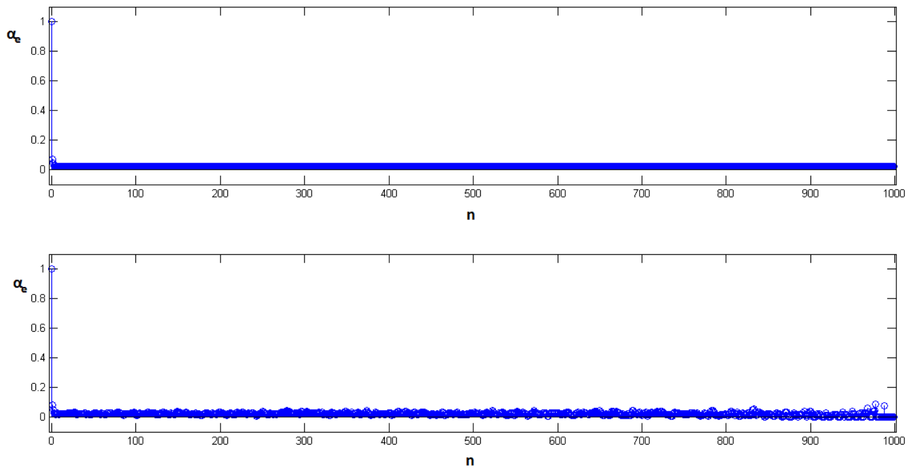
(a) The first subfigure of (a) illustrates the results from the theoretical derivation of the ACF whereas the second subfigure of (a) illustrates the results from the numerical estimation of the ACF.



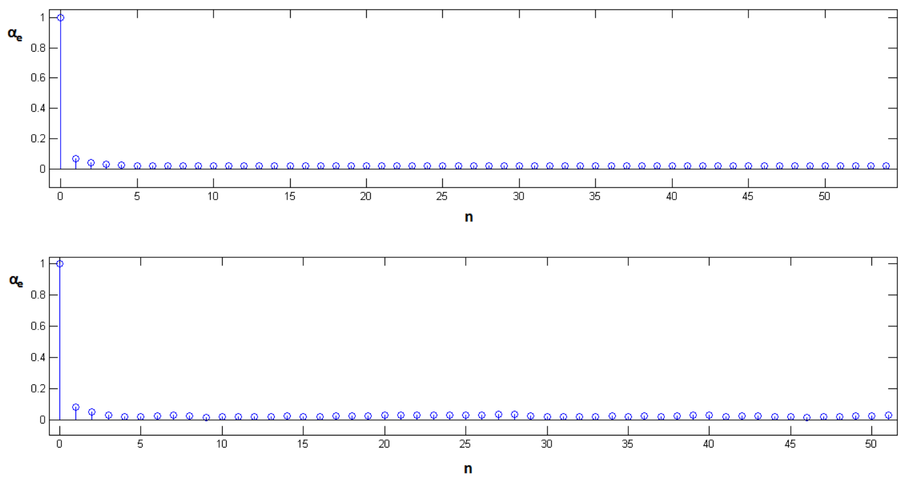
(b) For the zoomed in illustration of the results the same order of illustration with case (a) is followed.

Figure 4.4

The second simulation scenario investigates time correlation when the static agents are far distant the one to the other. An illustration of the obtained results is provided in figure 4.5.



(a) The first subfigure of (a) illustrates the results from the theoretical derivation of the ACF whereas the second subfigure of (a) illustrates the results from the numerical estimation of the ACF.



(b) For the zoomed in illustration of the results the same order of illustration with case (a) is followed.

Figure 4.5

Based on the obtained results in both cases the distribution becomes stationary over time. That type of outcome meets the results of the theoretical derivation of the ACF. In cases where the probability of having a connection is high correlation the distribution over a prolonged time period whereas for the interpolated distance in between the agents is long the distribution becomes stationary in a short time period.

That type of outcomes meet the behaviour of connectivity over a long time period. For the first case, where the agents are closely spaced with each other, it is expected that the agents will remain consecutively connected over long time periods until the connection drops. Thus the distribution needs a longer time period until it becomes independent on time. On the other hand in cases where the two agents are relatively far the one to the other the agents will remain connected over short time periods followed by longer time periods with no connection. Hence the distribution takes a shorter path to time independence.

4.2 Analysis of the ACF

The preliminary analysis upon the behaviour of the ACF over time indicated that there exist different cases regarding the time required for the distribution to become stationary. Another issue though that needs to be answered is what type of behaviour characterises the ACF until it yields a flat response over time. The stationary distribution consists of two parameters with an infinite pair of options for the same distribution. Thus we seek to investigate the variation of the outcome of the ACF for different pair of parameters. Therefore a step by step analysis is conducted starting with the most characteristic cases.

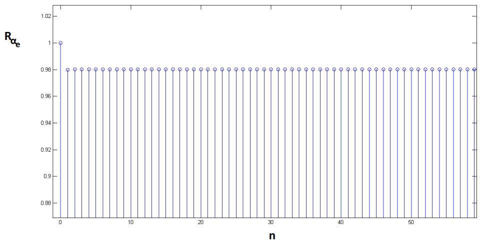
The main cases are separated into two groups depending on the probability of having a connection:

- A case where the probability for having a connection is high.
- A case where the connectivity is more likely to fail.

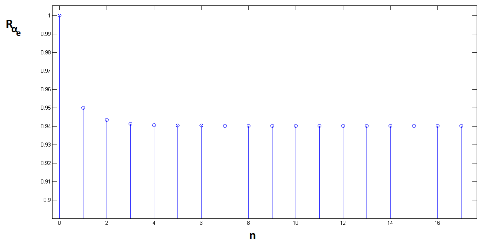
The preliminary analysis indicated that the first group is mainly characterised by three types of behaviour with respect to the convergence of the ACF:

- A flat response of the ACF.
- A case where the ACF converges fast.
- Smooth response of the ACF yields over time.

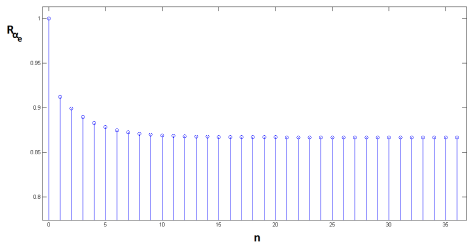
The two groups have common characteristics. The only difference for the second group there exists a case where the ACF fluctuates over time before converging to the steady state. In order to obtain a visual inspection upon the characteristic cases of the AFC, the different scenarios are illustrated in figure 4.6 . For the needs of the simulations we are considering the static case connectivity scenario all over the section. The reason why we follow this simulation scenario has do with the fact that the main interest upon the analysis of the ACF is to investigate the affect on connectivity which is related strictly on time.



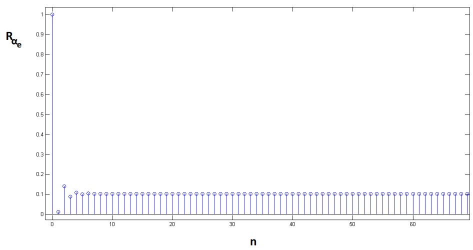
(a) flat response of the ACF



(b) Fast convergence of the ACF



(c) Smooth response of the ACF



(d) Fluctuation of the ACF

Figure 4.6

The next question that needs to be answered from the ACF analysis is what are the main factors that affect the ratio of the ACF convergence to the steady state. In order to proceed with this type of analysis a basic choice is made: The focus of the analysis is made upon the behaviour of the convergence ratio over the first time instance expressed by the following equation:

$$\mathbb{E}(a_{n+1} = 1 | a_n = 1) = \frac{\beta}{\beta + \alpha} \cdot (1 - \alpha). \tag{4.25}$$

The following time instances are expected to follow a behaviour based on the first time instance. In order to simplify our work this type of investigation is called first ACF coefficient analysis and the output of the corresponding function introduced in (4.25) is denoted by c. The behaviour of the output is depended to the form of the inputs which in this case are α, β. Based on this observations the first order component analysis can then be expressed in a form of function as follows:

$$[0,1] \times [0,1] \to [0,1]. \tag{4.26}$$

There exists an exceptional case for $(\alpha, \beta) = (0,0)$ where $c = \infty$. This scenario though is not considered since that case corresponds to 0–th instance of the transition matrix where no time evolution takes place. In order to obtain a visual inspection regarding the behaviour of $R_{a_e}(1)$ over the different pairs of (α, β) a plot involving the different coefficients of the first order component analysis is illustrated in figure 4.7.

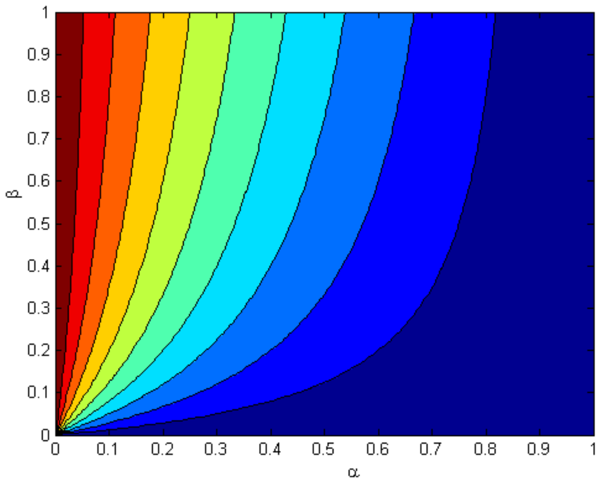


Figure 4.7: First ACF coefficient analysis over the region of α and β.

The obtained results indicate that the output of the corresponding result is mainly affected by the value of α. This tendency occurs because of the factor $1 - \alpha$ which has the major impact on the value of $R_{a_e}(1)$. There exist two extreme cases for $R_{a_e}(1)$. The first case holds for $R_{a_e}(1) = 1$ where in such case $\alpha =$

0. Considering the nature of the problem this means that once a connection in between the two agents is established the same connectivity status remains all over the time period regardless the value of β. An example of this scenario is illustrated in figure 4.8.

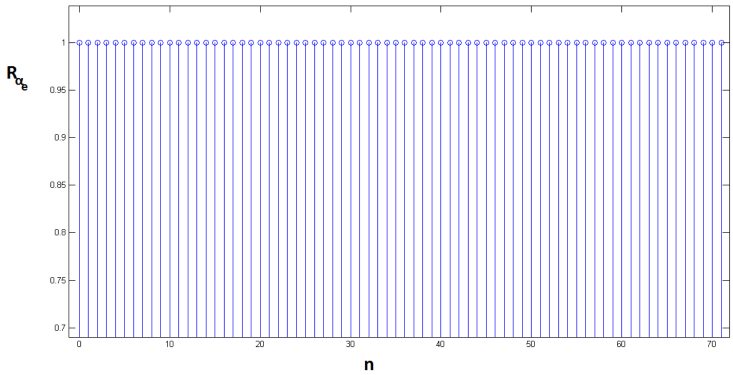


Figure 4.8: Flat response of the ACF where the two nodes remain connected all over the time period. The x axis represents time whereas the y axis corresponds to the value of c.

The second case holds for the case where $R_{a_e}(1) = 0$. There exist two possibilities for the current case. The first possibility holds for $\alpha = 1$. In such case the ACF is initially characterised by a fluctuation over time. The main characteristic of the current scenario is the fact that in case where a connection is established, there is not any case of remaining on the same status the next time instance. An example of this scenario is illustrated in figure 4.9. The second possibility holds when $\beta = 0$. In such case when the connectivity status switches to a state where no connection is established then the two agents loose connection permanently. The graphical representation of the ACF for the current possibility yields a flat zero response all over the time period.

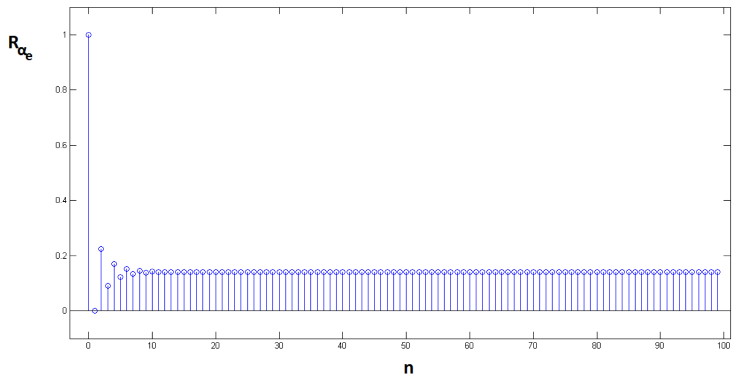


Figure 4.9: A scenario where $c = 0$ consequently the ACF is characterised by a fluctuation.The x axis represents time whereas the y axis corresponds to the value of c.

Once the basis of the analysis is defined and the main characteristics of the first ACF coefficient analysis are specified, the next step is to investigate the convergence ratio over the time period. Therefore an analysis upon this issue is conducted to investigate the factors that mainly affect the time elapsed for the ACF in order to reach the steady state. The main scenarios that are considered on the analysis correspond to the extreme cases:

- A case where the correlation between a state and the one that preceded it is high.
- A case where two consecutive time instances are weakly correlated.

For all the different cases all the coefficients of the two state *Markov* model are taken into account. The corresponding values are demonstrated in tables whereas a figure illustrating the convergence of the ACF is also provided.

The first simulation scenario considers the case where $R_{a_e}(1) = 0.92$. Two different forms of the coefficients are illustrated. In both cases the form of the transition matrix as well as the form of the stationary distribution are arbitrary chosen. The common aspect in both cases is the value of $R_{a_e}(1)$. The corresponding values of the components for the first case are illustrated in table 4.1.

Table 4.1: high value of $R_{a_e}(1) \to$ slow convergence

π_0	π_1	α	β	P_{00}
0.9515	0.0485	0.0331	0.5786	0.9669

The particular form of the coefficients yield the form of the convergence of the ACF illustrated in figure 4.10. As it can be seen from the results, the corresponding ratio of convergence for the particular form is relatively high. The ratio indicates a sudden drop from the first time instance to the second and smooth but short path to the steady state follows.

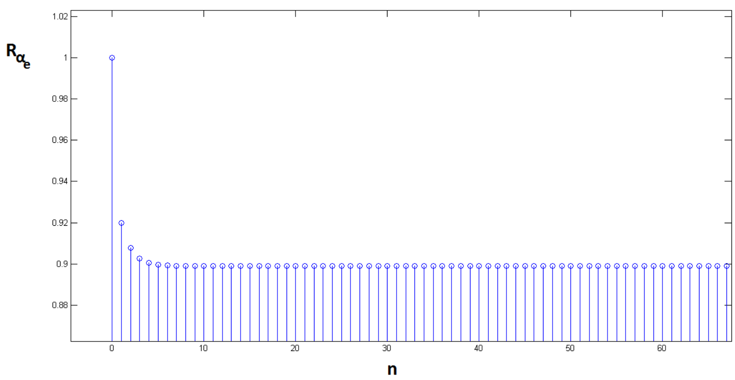


Figure 4.10

The analysis proceeds with another case for the same value of $R_{a_e}(1)$. The same process is followed to form the components of the two state *Markov* model and the obtained values are demonstrated in table 4.2.

Table 4.2: high $R_{a_e}(1) \rightarrow$ slower convergence

π_0	π_1	α	β	P_{00}
0.9379	0.0621	0.0191	0.2888	0.9809

The graphical representation of the current form of the *Markov* model is provided in figure 4.11. Unlike the first case, the current form of the *Markov* model introduces a smooth pattern to the ACF convergence ratio as well as a longer time period is required to reach the steady state.

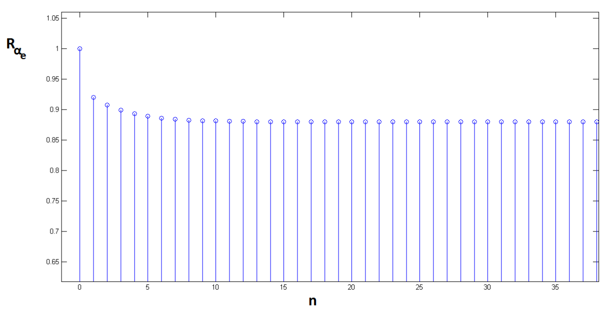


Figure 4.11

The second case,considering the level of correlation, investigates a simulation scenario where $R_{a_e}(1) = 0.002$. The first example where the components of the

Markov model meet this requirement is illustrated in table 4.3.

Table 4.3: low convergence component fast convergence

π_0	π_1	α	β	P_{00}
0.1129	0.8871	0.9823	0.1250	0.0177

The particular form of the coefficients yields the graphical representation illustrated in figure 4.12. As it can be seen from the obtained results the ratio is characterised by a sudden drop and a very small fluctuation over the first two time slots. The general pattern indicates a fast convergence to the steady state.

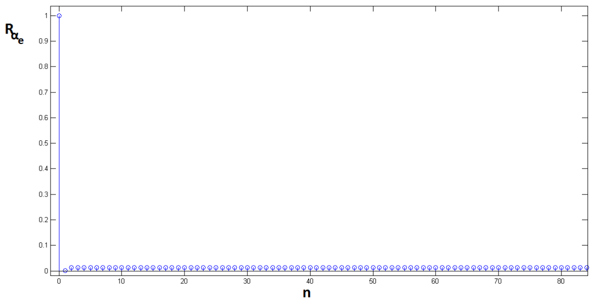


Figure 4.12

Another case for the same value of c is considered. The corresponding form of the components is introduced in table 4.4.

Table 4.4: low convergence component slow convergence

π_0	π_1	α	β	P_{00}
0.3056	0.6944	0.9935	0.4371	0.0065

The graphical representation that corresponds to the current form of the *Markovian* model is demonstrated in figure 4.13. As it can be seen from the obtained results, the current form of the coefficients introduces a higher fluctuation of convergence than before which yields a longer path to reach the steady state.

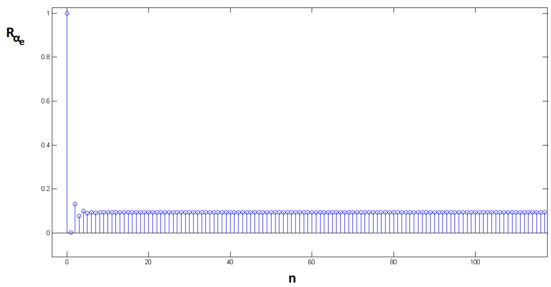


Figure 4.13

4.2.1 Conclusion

The current section provided an analysis on the ACF of the two state *Markov* model. In order to do so, a basic choice is made: The analysis is conducted with the focus on the first coefficient of the ACF. The obtained results indicated that there exists a variation on the convergence ratio fo the ACF for the same level of correlation. That kind of behaviour though is expected. The basic factor that affects the convergence ratio is the first component of the transition matrix. In cases where P_{00} is very close to π_0 the convergence to the steady state is very fast or even flat. On the other hand in cases where a long time period is required for P_{00} to converge to π_0, the corresponding pattern for the convergence ratio of the ACF is prolonged and in most cases smooth. The selection upon the entries of P is an important issue for the *Markovian* model. The connectivity in between a pair of agents is expected to have a time dependency that decays over time until the distribution becomes stationary thus the connectivity in between the agents is no longer time dependent. For the same outcome of the first ACF coefficient an infinite amount of pairs (α, β) can be selected. Considering the a nature of the connectivity issue it is preferable to select a pair of entries that introduces a smooth pattern until ACF converges. The analysis indicated that the entry that mainly affects the outcome of the first ACF component is α. Thus one way to address the issue of selecting the pair of entries could be by setting α such that, along with outcome of the first ACF component, the pair of entries meet the properties of the connectivity issue.

4.3 ACF coefficients analysis

One of the mains aspects of the connectivity problem is the correlation among the different time instances. Therefore an analysis of the *Markovian* model is provided. To do so a basic choice is made: The focus of the analysis is made with respect to the first component of the function denoted as $R_a(1)$. The conducted analysis on the behaviour of the ACF indicated that the convergence of R_a relies on the form of P_{00}. In order to obtain a deeper idea regarding the overall behaviour of R_a over the different components of the function a more thorough investigation is required. Thus an analysis is conducted to investigate the relationship among the different components. Since R_a consists of an infinite number of coefficients the investigation focus on the relevance among characteristic coefficients of ACF with respect to the first component. In particular the relevance is considered as having a repetition of the same pattern among the closest component to $R_a(1)$, in that case $R_a(2)$, and the more distant component,in that case $R_a(\infty)$. The cases are as follows:

- $R_a(1) = R_a(2)$.
- $R_a(1) = R_a(\infty)$.

Considering the first scenario among the different coefficients of R_a, the current relationship can be rewritten as follows:

$$\frac{\beta}{\beta+\alpha} \cdot (1-\alpha) = \frac{\beta}{\beta+\alpha} \cdot (\beta \cdot \alpha + (1-\alpha)^2). \tag{4.27}$$

The equation introduced in (4.27) can be rewritten in a one sided factorization equation form as follows:

$$\frac{\beta}{\beta+\alpha} \cdot (1-\alpha-\beta \cdot \alpha - (1-\alpha)^2) = 0. \tag{4.28}$$

The solutions of the current equation correspond to the case where the pattern of $\mathbb{R}_a$ introduced at the first time instance repeats in the second. The cases are as follows:

- $\alpha = 0 \Rightarrow \pi_0 = 1$
- $\beta = 0 \Rightarrow \pi_0 = 0$
- $\alpha + \beta = 1 \Rightarrow \pi_0 = \beta$

By following the same methodology as the one introduced for the first case, the second scenario can be expressed as follows:

$$\frac{\beta}{\beta+\alpha} \cdot (1-\alpha) = \lim_{n\to\infty} \frac{\beta}{\beta+\alpha} \cdot (\frac{\beta}{\beta+\alpha} + \frac{\alpha}{\beta+\alpha}(1-\alpha-\beta)^n) = (\frac{\beta}{\beta+\alpha})^2. \tag{4.29}$$

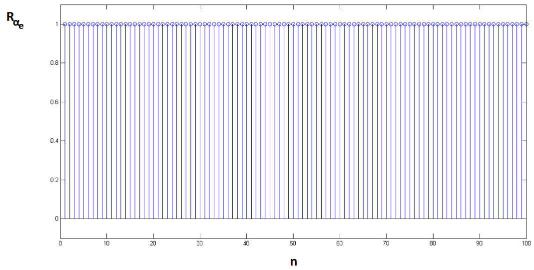
The current relationship can be expressed as a one sided factorised relationship as follows:

$$\frac{\beta}{\beta+\alpha} \cdot (1-\alpha-\frac{\beta}{\beta+\alpha}) = 0. \tag{4.30}$$

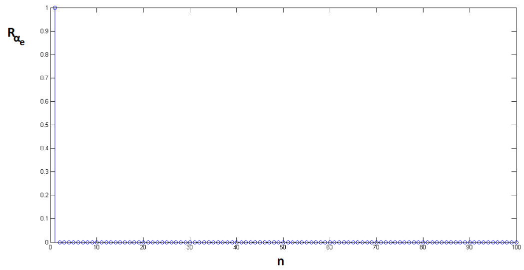
The scenarios where the same pattern introduced for where $\mathbb{R}_a(1)$ follows the same behaviour after after a very long time period holds for the following cases:

- $\alpha = 0 \Rightarrow \pi_0 = 1.$
- $\beta = 0 \Rightarrow \pi_0 = 0.$
- $\alpha + \beta = 1 \Rightarrow \pi_0 = \beta$

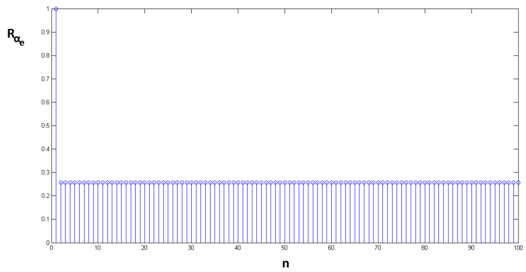
. For both cases the solutions are identical. That type of result is expected for the first two solutions , since in both cases $\mathbb{R}_a$ yields a flat response over the different time instances where for the first solution the agents are connected permanently whereas the opposite scenario holds for the second solution. The last solution corresponds to the scenario where the response of $\mathbb{R}_a$ is flat over time but the connectivity status is not constant. This kind of scenario holds for the case where rows of the transition matrix are identical or else $P_{11} = \pi_0$. In such case P is independent on time from the very first time instance. An illustration of the three different cases is provided in figure 4.14.



(a) Flat 1 response of the ACF→Infinite duration of connectivity.



(b) Flat 0 response of the ACF→Permanent loss of connection.



(c) Flat response with an ambiguity regarding the connectivity status.

Figure 4.14

The common aspect of the three cases, is that the *Markovian* model is independent on time thus it neglects to the model of *Savic* and *Zazo*. Derived on this we can claim that the *Markovian* model is a generalization of the model of *Savic* and *Zazo*. To verify this conclusion simple simulation scenario is conducted. The connectivity status for the model of *Savic* and *Zazo* is defined with respect to the corresponding probabilities. In case of the *Markovian* model the obtained probabilities from the model of *Savic* and *Zazo* are treated as the stationary distribution. Since the transition matrix is independent on time then the stationary distribution forms the rows of P. In order to provide a fair comparison the probabilities for both models are compared with a random value drawn from a uniform distribution which is common for the two models for the same time instance. The components of the *Markovian* model drawn from an arbitrarily chosen distribution are introduced in table 4.5.

Table 4.5: flat response of an arbitrary scenario where $\alpha + \beta = 1$.

π_0	π_1	α	β
0.1353	0.8647	0.1353	0.8647

The obtained results are illustrated in figure 4.15. As it can be seen from the results the two models introduce identical behaviour as expected. The same behaviour is expected to be obtained also from the two other solutions. For $\alpha = 0$ the two models will yield a permanent connection for the two agents all over the time period whereas for $\beta = 0$ then $\alpha = 1$ thus even of there was a connection at time $t = 0$, connection drops at time $t = 1$, and remains at a status where no connection exists permanently.

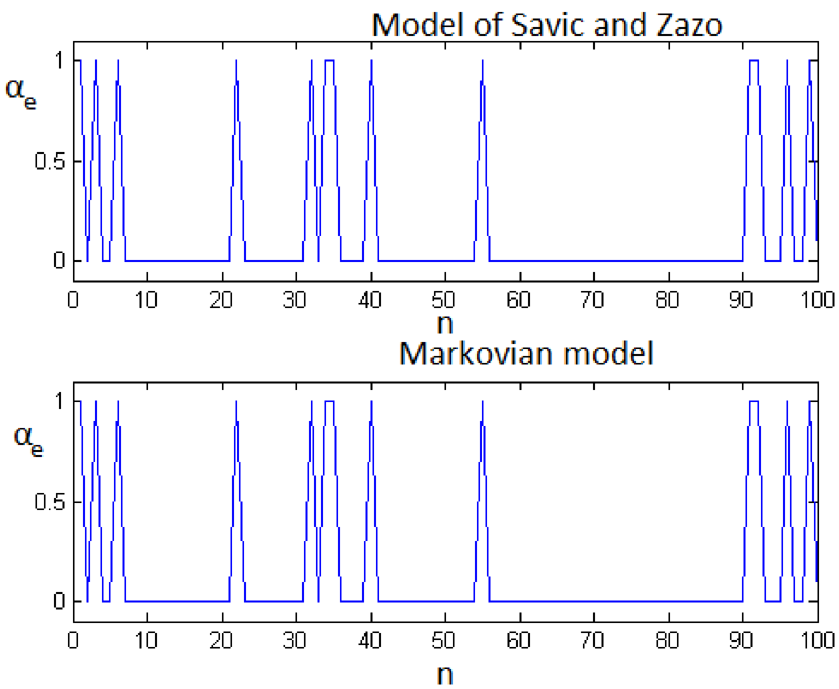


Figure 4.15: Comparison of the two models when R_a yields a flat response. The x axis represents the time instances and the y axis represents the connectivity status.

4.3.1 Pairwise selection of α, β based on the ACF coefficients analysis

The analysis upon the characteristic moments of the ACF indicated that there exist three basic scenarios upon the relationship in between the entries of the stochastic matrix that lead to a flat output of the ACF. The next question that needs to be answered is what is the affect on the output of the ACF when the pair of values exceeds the area formed by the interconnection of these points. A visual inspection of the interconnection region is provided in figure 4.16.

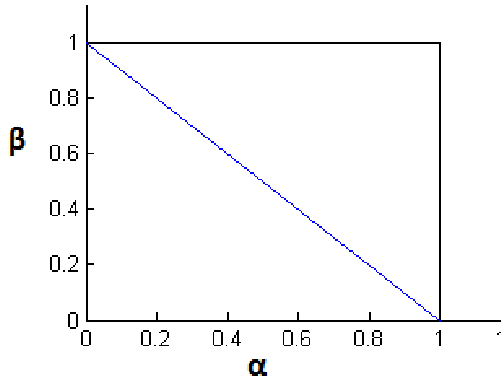


Figure 4.16: The diagonal of the rectangular region separates the area where the pair of entries of P yields a fluctuation to the ACF compared to the area that takes the form of a lower triangle and a where the pair of values yields a smooth pattern to the ACF.

The geometrical outcome of the interconnection region forms a triangle that splits the area upon the pair of values can be selected in half. A new question that needs to be answered due to the particular outcome is what are the differences in between the selection of a pair of entries located in the upper triangle compared to a pair of values located in the lower triangle. In practise this leads to a situation where the sum along the inverse diagonal, where the coefficients of the matrix which lead to the next are located, is larger than the sum of the main diagonal. The outcome of this situation is that the connectivity status is more likely to transit to the next state rather than staying at the current state. As a result the output of the ACF is characterised by a fluctuation which means that for the particular pair of values the model does not meet the *Markov* property. An example of such an outcome is illustrated in figure 4.13. Derived on this we can deduce that one of the basic requirements that need to be fulfilled by the pair of entries so that the model meets the *Markov* property is the following:

$$\alpha + \beta \leq 1. \tag{4.31}$$

4.4 Linear relationship of the components of P

The *Markovian* model introduces a generalization of the model proposed by *Savic* and *Zazo*. The current approach states the idea that the probability for having a connection obtained by the initial model corresponds to a stationary distribution. In such case the probability for having a connection the current state depends also on the state that preceded it. The relationship in between the two approaches can be expressed as follows:

$$\pi_0 = e^{\frac{-|x_1 - x_2|^2}{2 \cdot R^2}} = \frac{\beta}{\beta + \alpha} \tag{4.32}$$

The original approach introduces a single parameter, denoted as R,that controls the corresponding probability thus characterizes the behaviour of the model. Treating the corresponding probability as a stationary distribution introduces two entries for P, α and β, in the design of the corresponding model. Thus the infrastructure of the *Markovian* model consists of three basic parameters which need to be specified when constructing the model.

The main advantage when considering the *Markovian* approach to address the connectivity issue is that a model characterised by more than one parameters introduces a higher degree of freedom to the corresponding design. Thus the current model provides a flexible approach when considering the settings of a possible application. On the other hand, the corresponding parameters are characterised by several limitations which need to be considered when designing the model.

4.4.1 Limitations upon the α and β coefficients

The entries α and β of P introduce a high degree of freedom upon the selection of the corresponding values. The main relationship in between α and β is expressed in equation (4.32). Assuming that the value of R, the equation introduced in (4.32) can be rewritten such that one of the entries ,i.e β, can be expressed as a function of the second component. The current relationship can be expressed as follows:

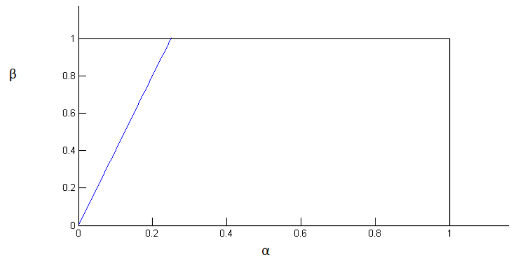
$$\frac{\beta}{\alpha + \beta} = \pi_0 \Rightarrow \beta = \pi_0 \cdot (\alpha + \beta) \Rightarrow \beta \cdot (1 - \pi_0) = \pi_0 \cdot \alpha \Rightarrow \beta = \frac{\pi_0 \cdot \alpha}{\pi_1} \tag{4.33}$$

The current equation expresses the linear relationship in between the components of the *Markovian* model. There exits an infinite combination of input and output values which could form the stationary distribution. Each of the possible outcomes though must fulfil the requirements of the stationary distribution introduced in 4.1. Derived on this there exist several cases that introduce a limitation upon the degree of freedom for each of the values of α and β. The limitation is introduced by the ratio $\frac{\pi_0 \cdot \alpha}{\pi_1}$. The different cases can then be listed as follows:

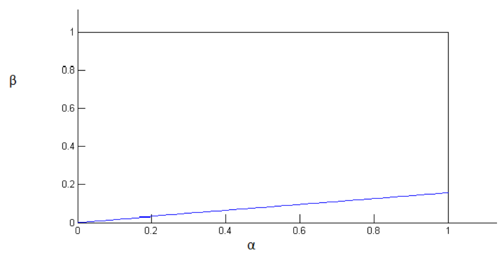
- $\frac{\pi_0}{\pi_1} < 1$, the limitation is introduced in the selection of β .

- $\frac{\pi_0}{\pi_1} > 1$, the limitation is introduced in the selection of α .
- $\frac{\pi_0}{\pi_1} = 1$, there is no limitation upon the selection of the components.

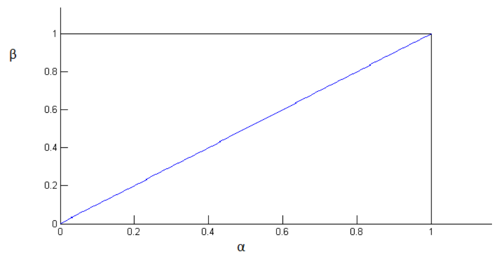
In order to obtain a visual inspection upon the possible outcomes an illustration of the different cases is provided in figure 4.17. The limitation upon the selection of the corresponding entry is introduced by the intersection of the equation introduced in (4.33) with the requirements of the stationary distribution.



(a) $\frac{\pi_0}{\pi_1} < 1$, thus the selection upon the selection of α is restricted.



(b) $\frac{\pi_0}{\pi_1} > 1$, thus the selection upon the value of β is restricted.



(c) $\frac{\pi_0}{\pi_1} = 1$, the selection of the components is free of limitations.

Figure 4.17: Illustration upon the different cases for the limitation upon the selection ofβ. The boxes illustrate the possible values for the components with respect to the requirements of the stationary distribution. The line illustrated in the figures corresponds to the input output relationship of the components for a fixed ratio $\frac{\pi_0}{\pi_1}$. The intersection in between the different aspects introduces the limitation upon the selection of the corresponding component.

The obtained results indicate that the design of a model where one of the entries, α and β, is constant is not a trivial task. In such case, i.e when α is constant, the main requirement that needs to be fulfilled is that $0 \leq \alpha \leq \alpha_{max}$, where α_{max} corresponds to the intersection point in between the rectangular area which represents the theoretical feasible region of the values and the representation of the linear relationship in between the values.

4.5 Model parameters

The *Markovian* model is characterised by a level of flexibility that allows to the model to become adoptive under different circumstances. One of the main aspects when considering the model is what kind of methods can be derived so we can select the proper model design. The linear relationship in between the entries of P indicates that the *Markovian* approach of the original model inserts one extra parameter to a possible design. Thus when considering a possible implementation of the model we are seeking for methods to calibrate a dual parametric model which for the *Markovian* model are denoted as R and α. The analysis upon the ACF coefficients along with the restrictions from the linear relationship in between the entries of P narrows down the region upon α can be selected so that the model meets the *Markov* property. By inserting the linear combination in between the entries of P expressed in equation (4.33) into the outcome from the AFC coefficients analysis expressed in (4.31) we can deduce that on of the minimum requirements that the α parameter should fulfil is the following:

$$\alpha + \beta \leq 1 \Leftrightarrow \alpha + \frac{\pi_0}{\pi_1} \cdot \alpha \leq 1 \Leftrightarrow \alpha(\pi_0 + \pi_1) \leq \pi_1 \Leftrightarrow \alpha \leq \pi_1 \tag{4.34}$$

No further information upon α can be obtained though by the so far conducted analysis. Hence in order to calibrate the α we need some information from the signal propagation environment upon the model may be implemented. On the other hand, there is not any need for the R parameter to fulfil any particular property. The calibration of R is mainly related with the signal propagation environment or it can be done in a theoretical level is comparison with another method that address the connectivity issue i.e the model of *Henk Wymeersch*.

4.5.1 Setting the α parameter

A possible implementation of the *Markovian* model in a specific region can be conducted with by having a set data available. One approach to set the α is in cases where the data set provides information regarding the maximum time length that a node was consecutively connected to the network. The theoretical expression considering a time length of k upon this relationship reads:

$$\sum_{k=0}^{\infty} k \cdot (1-\alpha)^k \cdot \alpha. \tag{4.35}$$

The total expectation is expressed of terms which form an infinite sequence. In order to obtain the expectation for the duration length the first step is to investigate whether this sequence converges or not. To do so we can utilize the *D'Alembert's* criterion which is expressed as follows:

$$\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} < 1 \tag{4.36}$$

By inserting the sequence which expresses the length of duration in the criterion the equation introduced in (4.36) can be rewritten as follows:

$$\lim_{k\to\infty} \frac{(k+1)\cdot(1-\alpha)^{k+1}\cdot\alpha}{k\cdot(1-\alpha)^k\cdot\alpha} = (1-\alpha)\cdot\lim_{k\to\infty}\frac{(k+1)}{k} = (1-\alpha) < 1, \tag{4.37}$$

thus the criterion is fulfilled.

Once the convergence of the series has been verified, the next step in our analysis is to compute the value where the sum converges. The first step of the current task is to provide the analytical expression of the series:

$$\sum_{k=0}^{\infty} k\cdot(1-\alpha)^k\cdot\alpha = (1-\alpha) + 2\cdot(1-\alpha)^2 + \ldots k\cdot(1-\alpha)^k. \tag{4.38}$$

.

There exists a common ratio r in between the terms of the sequence expressed as follows:

$$r = \frac{(k+1)}{k}\cdot(1-\alpha). \tag{4.39}$$

.

The form of the series can be interpreted as geometric series.A very common approach to express geometric series is as it follows :

$$S_k = d\cdot\frac{1-r^k}{1-r}, \text{where } d = 1-\alpha \text{ for the current problem.} \tag{4.40}$$

.

Derived on this and with respect to the current form of r and d, S_k can be expressed as follows:

$$S_k = (1-\alpha)\cdot\frac{1-(\frac{(k+1)}{k}\cdot(1-\alpha))^k}{1-\frac{(k+1)}{k}\cdot(1-\alpha)}. \tag{4.41}$$

.

Since we are seeking for behaviour of S_k while $k \to \infty$ the main focus is upon the terms of the equation related to k. By inserting the result from equation (4.37), equation (4.41) can be rewritten as follows:

$$\lim_{k\to\infty} S_k = (1-\alpha) \cdot (\frac{1-(1-\alpha)^k}{1-(1-\alpha)}). \tag{4.42}$$

.

The only term that depends on k is $(1-\alpha) < 1$ which goes to 0 when $k \to \infty$, thus the final form of the equation can be expressed as follows:

$$\lim_{k\to\infty} S_k = \frac{1-\alpha}{\alpha} \tag{4.43}$$

.

4.6 Connectivity of a mobile agent with neighbouring base nodes

The so far conducted research of the current work,considering the different connectivity models, focus on the main characteristics of each model and an analysis is provided regarding the advantages and disadvantages of each approach. One of the main deficiencies when attempting to conduct a deeper analysis arises from the act that the so far developed approaches are structured in a way such that no straight forward comparison can be made on a general basis. To illustrate this argument we introduce the behaviour of the two models on a topological basis. So far we investigated the behaviour of the models with respect to a network with the minimum amount of participants, the current section investigates the behaviour of of the model when multiple agents participate in the network. In our research we consider the simplest case where a mobile agent communicate with two anchor nodes. An extension of the conclusion based on the obtained results is expected to be straight forward when considering larger networks.

4.6.1 The model of *Savic* and *Zazo*

The model proposed by *Savic* and *Zazo* exploits the probabilistic variation of having a connection depending on distance. A main characteristic of the current approach is the fact that the established connections are independent with each other. Thus when considering a scenario where a mobile agent abbreviated as 1 attempts to connect with two anchor nodes abbreviated as 2 and 3 respectively the corresponding probabilities can be expressed as follows:

$$\mathbb{E}[a_{12}(n)] = \mathbb{P}[a_{12}(n) = 1|x_1, x_2] = e^{\frac{-|x_1-x_2|^2}{2\cdot R^2}} \tag{4.44}$$

and

$$\mathbb{E}[a_{13}(n)] = \mathbb{P}[a_{13}(n) = 1|x_1, x_3] = e^{\frac{-|x_1 - x_3|^2}{2 \cdot R^2}} \tag{4.45}$$

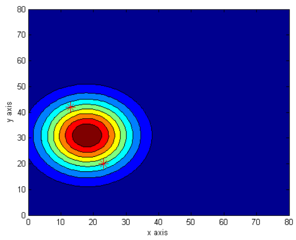
Derived on these, the probability for the mobile agent to have a connection with both the other two agents can be expressed as follows:

$$\mathbb{E}[a_{12}(n) \cdot a_{13}(n)] = \mathbb{P}[a_{12}(n) = 1|(x_1, x_2), a_{13}(n) = 1|x_1, x_3] \tag{4.46}$$

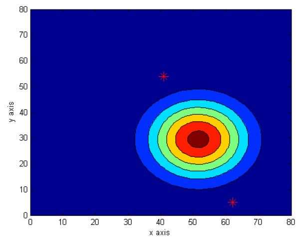
Since the two incidents are independent with each other the equation introduced in (4.46) can be rewritten as follows:

$$\mathbb{E}[a_{12}(n)] \cdot \mathbb{E}[a_{13}(n)] = e^{\frac{-(|x_1 - x_2|^2 + |x_1 - x_3|^2)}{2 \cdot R^2}} \tag{4.47}$$

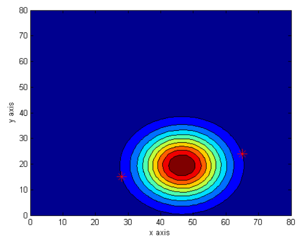
In order to obtain a visual inspection upon the current relationship, a simple experiment is conducted. A mobile agent travels around a predefined simulation area whereas the two other nodes are considered to be anchor nodes thus their location remains static over time. The corresponding probabilities are obtained with respect to the model of *Savic* and *Zazo* and the obtained results are illustrated in figure 4.18.



(a) A case where the Diameter of the output exceeds the distance in between the base nodes



(b) A case where the Diameter of the output is shorter than the distance in between the base nodes



(c) A case where the Diameter of the circular output equals the distance in between the base nodes

Figure 4.18: A contour plot of the function introduced in Equation (4.46). The x and y axis represents correspond to the coordinates of the agents. The location of the base nodes is marked with a red cross.

As it can be seen from the illustrated results, the output of the function introduced in Equation (4.46) takes the form of a circular *Gaussian* PDF. The particular form of the output is expected since the form of the function is similar to the equation that corresponds to the *Normal* distribution. Based on the results we can deduce that there exist three characteristic cases for the output of the function:

- The diameter of the circle exceeds the distance D in between the base nodes.
- The diameter of the circle equals the distance in between the base nodes.
- The diameter of the circle is shorter than the distance in between the base nodes.

The main factor that affects the output of the function is the parameter of the model of *Savic* and *Zazo* denoted as $\mathbb{R}$. Thus one of the main considerations when using such model is how to define the parameter in a way that fits the requirements of the implemented application.

4.6.2 The Connectivity model of *Henk Wymeersch*

The connectivity model proposed by *Henk Wymmersch* introduces a fairly simple approach to define of either a connection in between two agents can be established or not, depending on distance. Considering the case where the mobile agent(denoted as 1) attempts to connect with two base nodes(denoted as 2 and 3 accordingly) where each of them has a given radius R_1 and R_2, then the mobile agent is connected with the base nodes if it is located in the circular surfaces A_1,A_2 where:

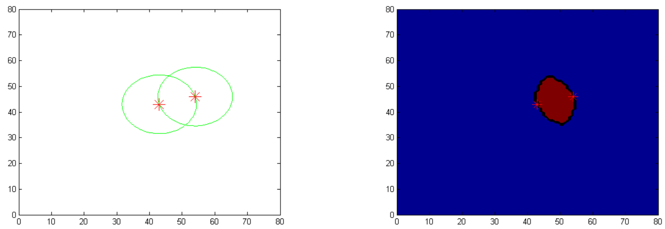
$$A_1 = [||x_1 - x_2||_2 < R_{H_1}] \text{ and } A_2 = [||x_1 - x_3||_2 < R_{H_2}] \tag{4.48}$$

Considering the case where the mobile agent attempts to connect with two base nodes simultaneously then the in order for this case to come true, we should consider that the corresponding distances are dependent since both of them depend on the location of x_1. Thus the case where 1 is connected with the first 2 cannot be investigated in separate with the case where 1 attempts to connect with the 3. Derived on this the corresponding function can be expressed as follows:

$$\mathbb{E}[a_{12}(n) \cdot a_{13}(n)] = A_1 \cap A_2 \tag{4.49}$$

.

In order to obtain a geometrical representation a simple experiment is conducted. Two base nodes are randomly placed on predefined simulation area with the same transmission radius which is equal to the distance between them. In the same time a mobile agent travels around all over the simulation area. The corresponding results are illustrated in figure 4.19.



(a) The location of the base nodes along with the transmission radius

(b) The area where the mobile agent is connected with the base nodes simultaneously.

Figure 4.19: An illustration of the case where a mobile agent is connected with two base nodes with respect to the connectivity model of *Henk Wymeersch*.Both base nodes have a transmission radius which is equal to the distance between them. The first subfigure illustrates the location of the base nodes along with the transmission radius whereas the second subfigure illustrates the area where the mobile agent is connected with the base nodes simultaneously.

The obtained results indicate the differences in between the two approaches on a topological level. The area of interest for the model of *Savic* and *Zazo* takes the from of a circular pdf whereas in case of the model of *Henk Wymmersch* it similar to a rugby ball. Thus the models cannot be compared on a topological level. The same argument holds for networks with more participants.

4.6.3 Setting the R parameter

The *Markovian* model is a generalization of the approach introduced by *Savic* and *Zazo*. The initial model is characterised by a single parameter, denoted as R, which can be interpreted as the signal power. A possible implementation of the *Markovian* model requires the specification of the particular component thus there is a need of a pattern upon this choice can be conducted. Ideally the model can be calibrated by making use of mass data. This case does nit hold for the current work though. Thus there is need to develop an alternative method upon such choice can be made. A common practise in such case is to calibrate the parameters of one model with respect to another approach followed in the same field. For the current case this can be done through the connectivity model of *Henk Wymeersch*.

There are several difficulties when attempting to set the range parameter of the model of *Savic* and *Zazo* with respect to the model of *Henk Wymeersch*. The main difficulty has to do with the fact that the two approaches are completely different. In the first case the proposal considers a probabilistic approach to address the issue while on the other hand the model of *Henk Wymeersch* is defined with respect to a deterministic function. In order to obtain a visual inspection upon the differences in between the different approaches an illustration of the output of the corresponding functions is provided in figure 4.20.

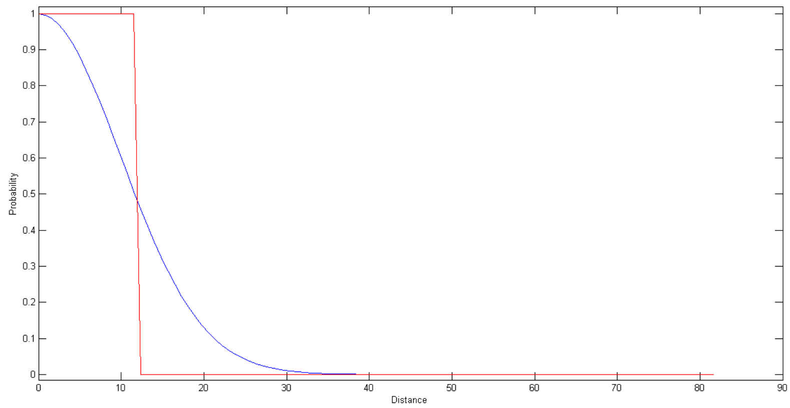


Figure 4.20: The x axis represents the distance in between the agents whereas the y axis represents the probability for having a connection.

As it can be seen from the results the output of the probabilistic model can be considered as a one sided *Gaussian* kernel while the output of the model of *Henk Wymeersch* takes the form of a rectangle. Thus a comparison in between the two approaches requires to find a way to fit the one shape to the other. Such task though is by default quite difficult.

One approach to calibrate the range parameter of the probabilistic model is set R such that the two methods to have a similar performance within the transmission radius. In such case R should be set in a way such that the probability for having a connection within the transmission radius is relatively very high. A simple approach to fulfil such task is to express R as a function of probability , denoted as ϵ, so we can tune the parameter of the probabilistic model of *Savic* and *Zazo* to is as close to R_H as we wish with respect to ϵ. Derived on this, R can be expressed in accordance with R_H by making use of the following probabilistic rule:

$$R(\epsilon) = \sqrt{-\frac{R_H^2}{2 \cdot ln\epsilon}}, \text{where } 0 < \epsilon < 1 \tag{4.50}$$

.

In order to obtain an visual inspection upon the behaviour of the probabilistic rule an illustration of a case where R is calibrated with respect to R_H for a specific value of ϵ. The obtained results are illustrated in figure 4.21.

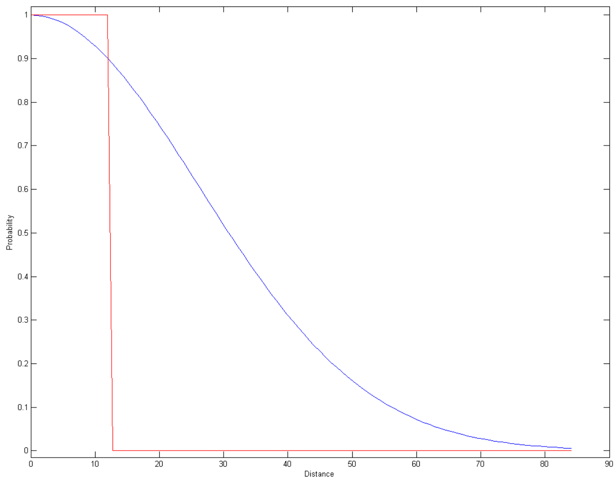


Figure 4.21: A static scenario where the range parameter is calibrated such that the two models introduce a similar performance.$\epsilon = 0.95$

As it can be seen from the results as long as the distance in between the agents is within the transmission radius the corresponding part of the one *Gaussian* function gets close to the upper side of the rectangular output of the model of . While the probability gets closer to 1 the surface included in between the output within the transmission radius shrinks. Thus one way to calibrate R with respect to R_H is to tune the model parameter such that it has a close performance with the model of *Henk Wymeersch* in a bounded area.

Chapter 5

Model Calibration

The so far developed work introduces a generalization of the original model proposed by *Savic* and *Zazo* in the time domain. A basic characteristic of the model is that it is based on multiple parameters thus it can be adopted by different set ups. The original model has a single parameter , R, whereas the generalization to the *Markovian* model that addresses the initial model as a stationary distribution adds an extra parameter in the overall set up.

A fair question when considering models, is what kind of strategy can be followed in order to select the corresponding parameters. In case of the *Markovian* model the parameters,α and β, of the stationary distribution are related with each other thus an estimation of one of the parameters,i.e α, reflects directly on the other. On the other hand the parameter ,R, of the initial model can be obtained in separate since it is not dependent on the rest of the parameters. Contrariwise an estimation of R reflects on the estimation of α.

Addressing the calibration of the model as an estimation problem we are seeking for methods to estimate the model parameters (α, R). The model parameters are estimated with respect to a given data set which in our case corresponds to the pairwise connectivity status $a_{e_ij}(n)$ for a particular time instance given the positions (x_i, x_j) of the agents of that moment. The dataset is for a single link is denoted as $A^{(d)} = (a_{e_{ij}}(n)^{(d)} | (x_i(n), x_j(n))^{(d)})$

There exist several criteria upon the corresponding methodologies can be developed. The most commonly used are the Mean–Square Error(MSE) and the Maximum a posterior Probability (MAP). When no prior information is available regarding the probability distribution of the estimated variable MAP neglects to the Maximum Likelihood(ML) estimation. The current work develops a mathematical framework upon the principle of the ML estimator in order to obtain estimates upon α and R. In case of R no prior information is considered so far thus the ML estimation is followed.

Two different approaches are developed to estimate the model parameters. The first approach develops the ML criterion for each of the parameters in separate. The R parameter is independent to α whereas the opposite case holds for α. Thus the first step in our method is to develop an ML criterion for R and then pass the information to the ML criterion for α. In case of a real life environment in order to develop requires in practise two different data sets. The first data set utilised to estimate R requires a data set where the available observations are independent on time. Thus the same type of measurements is repeated in between long time periods. On the other hand the ML criterion for α exploits the available information where the available measurements are obtained over consecutive time instances. The second approach address the estimation problem jointly for both the parameters. Thus given a data set an ML criterion is developed that estimates jointly the model parameters.

5.0.1 ML Estimator of R

The R parameter states the idea that the probability for having a connection relies on additive aspects, rather than distance, that affect the connectivity status in between the agents. Such an aspect could be considered for example the transmission power. Assuming that R expresses the power of the transmitted signal a basic choice that can be made for the nodes participating in a graph is that the transmission power for a particular time instance is common for all the agents. In such case the probability for several nodes to exchange information with each other varies depending on the distance in between them. Thus given the distances in between the agents of the graph the output of the connectivity status results a sequence of data,denoted as $A^{(d)}$, that specify whether a connection can be established in between the agents or not. Derived on these the Likelihood function reads:

$$\Lambda(R|A^{(d)}) = P(a_{e_{ij}} = a_{e_{ij}}^{(d)}|(x_i, x_j)^{(d)}, R) \tag{5.1}$$

The model proposed by *Savic* and *Zazo* reads:

$$P(a_{e_{ij}} = a_{e_{ij}}^{(d)}|(x_i, x_j)^{(d)}, R) = \begin{cases} e^{\frac{-|x_i - x_j|^2}{2 \cdot R^2}}, \text{if } a_{e_{ij}}^{(d)} = 1 \\ 1 - e^{\frac{-|x_i - x_j|^2}{2 \cdot R^2}}, \text{if } a_{e_{ij}}^{(d)} = 0 \end{cases} \tag{5.2}$$

The connections in between the agents are considered to be independent with each other, thus for a given data set $A^{(d)}$ with respect to equation introduced in (5.2) the likelihood function for R can be rewritten as follows:

$$\begin{aligned} \Lambda(R|A^{(d)}) = P(a_{e_{ij}} = a_{e_{ij}}^{(d)}|(x_i, x_j)^{(d)}, R) &= a_{e_{ij}}^{(d)}|(x_i, x_j)^{(d)}, R) \\ &= \prod_{a_{e_{ij}}^{(d)}=1} e^{\frac{-|x_i - x_j|^2}{2 \cdot R^2}} \prod_{a_{e_{ij}}^{(d)}=0} 1 - e^{\frac{-|x_i - x_j|^2}{2 \cdot R^2}} \end{aligned}$$

By making use of the properties of the exponential function the log likelihood function reads:

$$ln\Lambda(R|A^{(d)}) = \sum_{a_{e_{ij}}^{(d)}=1} \frac{-|x_i - x_j|^2}{2 \cdot R^2} + \sum_{a_{e_{ij}}^{(d)}=0} 1 - e^{\frac{-|x_i - x_j|^2}{2 \cdot R^2}}. \tag{5.3}$$

The estimate for R corresponds to the point point that maximizes the likelihood function. Derived on this the estimate for the parameter reads:

$$\hat{R}_{ML} = \arg\max_{R} \; ln\Lambda(R|A^{(d)}) \tag{5.4}$$

We are seeking to maximize the output of the function introduced in (5.3). A very common approach to accomplish such task is to compute the first order derivative of the corresponding function. Derived on this another estimate of R can be obtained as a solution of the following equation.

$$\frac{\partial ln\Lambda(R|A^{(d)})}{\partial R} = \text{argmax} \; \frac{1}{R^3}\Big(\sum_{a_{e_{ij}}^{(d)}=1} |x_i - x_j|^2 - \sum_{a_{e_{ij}}^{(d)}=0} \frac{e^{\frac{-|x_i - x_j|^2}{2 \cdot R^2}}}{1 - e^{\frac{-|x_i - x_j|^2}{2 \cdot R^2}}} \cdot |x_i - x_j|^2\Big) = 0 \tag{5.5}$$

Note that there exist root finder algorithms that can be utilized in alternative. In such case, methods such as the *Gauss–Newton* are applied to track the extreme of the function. For the current case one extra step is required which os the computation of the second order derivative of the ML function. The root finder algorithm tracks the extreme of the function when the first order derivative does not have many zero crossings. Another requirement is that the initial point must be very close to the solution. Thus there is need a strategy upon the initial point can be obtained. Hence the ML criterion could be used as strategy to derive the initial point. Consequently, in our case, the root finder algorithm relies on the performance of the ML criterion. The same argument holds for all the different cases that we develop the ML criterion.

5.0.2 ML Estimator of α

The estimation of the model parameters is a stepwise process in which the estimation of one parameter relies on the estimation of the other. R parameter can be estimated separately since the connectivity status relies upon distance whereas the estimation of the parameters of the stationary distribution depend on $\hat{R}$. Thus estimation of α follows the estimation of R and the performance of the corresponding estimator is affected by the performance of the estimator that preceded it. Derived on this the likelihood function of α reads:

The *Markovian* approach of the connectivity issue introduces a dependency in between the current status of connection and the state that preceded it. Based on this property and with respect to the form of the stationary distribution the different time instances can be categorised in sets as follows:

1. A=$\{$ α, if $a_{e_{ij}}(n) = 0|a_{e_{ij}}(n-1) = 1$ $\}$
2. B=$\{$ $\alpha \cdot \frac{\pi_0(n)}{\pi_1(n)}$, if $a_{e_{ij}}(n) = 1|a_{e_{ij}}(n)(n-1) = 0$ $\}$
3. Γ=$\{$ $1 - \alpha \cdot \frac{\pi_0(n)}{\pi_1(n)}$, if $a_{e_{ij}}(n) = 0|a_{e_{ij}}(n-1) = 0$
4. Δ=$\{$ $1 - \alpha$, if $a_{e_{ij}}(n) = 1|a_{e_{ij}}(n-1) = 1$ $\}$

The static nodes connectivity scenario

The basic characteristic when considering the estimation of α is that unlike $\hat{R}$ a time evolution is involved. In such case the positions of the node evolve over time consequently $\frac{\pi_0(n)}{\pi_1(n)}$ varies over time. The current subsection considers the case where the nodes remain static over time thus $\frac{\pi_0(n)}{\pi_1(n)}$ is unaltered over the different instances.

At any given time instance $\frac{\pi_0(n)}{\pi_1(n)}$ relies on $\hat{R}$ as well as the distance d_{ij} in between the agents. Thus with respect to the model proposed by *Savic* and *Zazo* $\frac{\pi_0(n)}{\pi_1(n)}$ can be obtained as follows: $(\frac{\pi_0(n)}{\pi_1(n)}|d_{ij)}, \hat{R})$. Since the node is considered to be static over time the n index drops. Derived on this and as soon as the estimation upon $\frac{\pi_0}{\pi_1}$ is obtained, the likelihood function for α when considering a single link reads:

$$\Lambda(\alpha|A^{(d)}, R) = \ s_0 \cdot \alpha^{|A|} \cdot (\alpha \cdot \frac{\pi_0}{\pi_1})^{|B|} \cdot (1 - \frac{\pi_0}{\pi_1} \cdot \alpha)^{|\Gamma|} \cdot (1-\alpha)^{|\Delta|} \tag{5.6}$$

where s_0 is defined as follows:

$$s_0 = \begin{cases} \pi_0(0), \text{if } a_{e_{ij}}^{(d)}(0) = 1 \\ \pi_1(0), \text{if } a_{e_{ij}}^{(d)}(0) = 0 \end{cases} \tag{5.7}$$

Thus the estimate for α reads is expressed as:

$$\hat{\alpha}_{ML} = arg \max_{\alpha} \ \Lambda(\alpha|A^{(d),R}) \tag{5.8}$$

Alternatively $\hat{\alpha}$ can be obtained as a solution of the following equation:

$$\begin{aligned}\frac{\partial\Lambda(\alpha|A^{(d)},R)}{\partial a} &= s_0(|A|\alpha^{|A|-1}\cdot(\alpha\cdot\frac{\pi_0}{\pi_1})^{|B|}\cdot(1-\frac{\pi_0}{\pi_1}\cdot\alpha)^{|\Gamma|}\cdot(1-\alpha)^{|\Delta|}\\ &+\alpha^{|A|}\cdot\frac{\pi_0}{\pi_1}\cdot|B|\cdot(\alpha\cdot\frac{\pi_0}{\pi_1})^{|B|-1}\cdot(1-\frac{\pi_0}{\pi_1}\cdot\alpha)^{|\Gamma|}\cdot(1-\alpha)^{|\Delta|}\\ &-\alpha^{|A|}\cdot(\alpha\cdot\frac{\pi_0}{\pi_1})^{|B|}\cdot|\Gamma|\frac{\pi_0}{\pi_1}\cdot(1-\frac{\pi_0}{\pi_1}\cdot\alpha)^{|\Gamma|-1}\cdot(1-\alpha)^{|\Delta|}\\ &-\alpha^{|A|}\cdot(\alpha\cdot\frac{\pi_0}{\pi_1})^{|B|}\cdot(1-\frac{\pi_0}{\pi_1}\cdot\alpha)^{|\Gamma|}\cdot|\Delta|\cdot(1-\alpha)^{|\Delta|-1})=0\end{aligned} \tag{5.9}$$

The dynamic scenario

The estimation of α when considering the static scenario can be extended for the case where the node is characterised by a dynamic behaviour regarding its location over time. In such case the stationary distribution varies over time thus the probabilities are estimated as follows:$(\frac{\pi_0(n)}{\pi_1(n)}|(x_i(n),x_j(n)),R)$. Derived on this the the likelihood function of α introduced in equation (5.6) can be rewritten as follows:

$$\Lambda(\alpha|A^{(d),R})=s_0\cdot\alpha^{|A|}\cdot(1-\alpha)^{|\Delta|}\cdot\prod_{n\in B}(\alpha\cdot\frac{\pi_0(n)}{\pi_1(n)})\cdot\prod_{n\in\Gamma}(1-\alpha\cdot\frac{\pi_0(n)}{\pi_1(n)}) \tag{5.10}$$

Hence estimate of α reads:

$$\hat{\alpha}_{ML}=arg\max_{\alpha}\ \Lambda(\alpha|A^{(d),R}) \tag{5.11}$$

Moreover $\hat{\alpha}$ can be obtained as a solution of the following equation:

$$\begin{aligned}\frac{\Lambda(\alpha|A^{(d),R})}{\partial a} &= s_0(|A|\cdot\alpha^{|A|-1}\cdot(1-\alpha)^{|\Delta|}\cdot\prod_{n\in B}(\alpha\cdot\frac{\pi_0(n)}{\pi_1(n)})\cdot\prod_{n\in\Gamma}(1-\alpha\cdot\frac{\pi_0(n)}{\pi_1(n)})\\ &+\alpha^{|A|}\cdot(1-\alpha)^{|\Delta|}\cdot|B|\alpha^{|B|-1}\cdot\prod_{n\in B}\frac{\pi_0(n)}{\pi_1(n)}\cdot\prod_{n\in\Gamma}(1-\alpha\cdot\frac{\pi_0(n)}{\pi_1(n)})\\ &+\alpha^{|A|}\cdot(1-\alpha)^{|\Delta|}\cdot\prod_{n\in B}(\alpha\cdot\frac{\pi_0(n)}{\pi_1(n)})\cdot(|\Gamma|\alpha^{|\Gamma|-1}\cdot\prod_{n\in\Gamma}(\frac{\pi_0(n)}{\pi_1(n)})-\sum_{n\in\Gamma}(\frac{\pi_0(n)}{\pi_1(n)}))\\ &-\alpha^{|A|}\cdot|\Delta|\cdot(1-\alpha)^{|\Delta|-1}\prod_{n\in B}(\alpha\cdot\frac{\pi_0(n)}{\pi_1(n)})\cdot\prod_{n\in\Gamma}(1-\alpha\cdot\frac{\pi_0(n)}{\pi_1(n)}))=0\end{aligned} \tag{5.12}$$

Simulation Results

The current subsection aims to provide an evaluation of the performance of the proposed estimators. Two different approaches are followed to fulfil such task. At first each an evaluation on the performance of the estimator for the R parameter is given since it is the one that is independent on the other parameters. In order to obtain a full idea of the performance estimators a simultaneous estimation upon the parameters α, R via a message passing algorithm.

The estimation of $\hat{R}$ is characterised by the fact that no time evolution is required in order to estimate the parameter. A number of measurements along with the corresponding connectivity status are sufficient to obtain an estimation upon $\hat{R}$. Derived on this a simulation scenario where a number of N measurements with the connectivity status is considered to be known. The corresponding algorithm provides an estimation upon $\hat{R}$ by searching for the value that maximizes ML in a specified region. The performance of the estimator is conducted with respect to the mean error(ME) criterion, where ME is as follows:

$$ME_{\hat{R}} = \frac{1}{N} \cdot \sum_{i=1}^{N} |\hat{R}(i) - R(i)| \tag{5.13}$$

The components of the simulation set up as well as the result for ME are provided in table 5.1.

Table 5.1: Simulation parameters and result

candidates	Search Region	Number of Measurements	ME
200	(8,35)	800	1.2

In order to provide a visual inspection upon the performance of the ML estimator a snapshot from the first 20 realizations is provided in figure 5.1.

The results indicate that the ML estimator is in general a good approach to estimate $\hat{R}$. There exist several cases though where the estimator is characterised by an extreme performance whether it is poor or accurate. In order to obtain a better idea why this phenomenon occurs an in depth analysis based on the available information is conducted. Since the available data are related to the connectivity status and the corresponding measurements. The comparison in between a god accuracy measurement of ML and the poorest performance of the criterion is illustrated in figure 5.2.

The obtained results indicate that the performance of the ML criterion is high when the output of the connectivity status yields a relatively sparse sequence output. In contrary when the connectivity over time is relatively dense the ML criterion fails to provide a good estimation for $\hat{R}$.

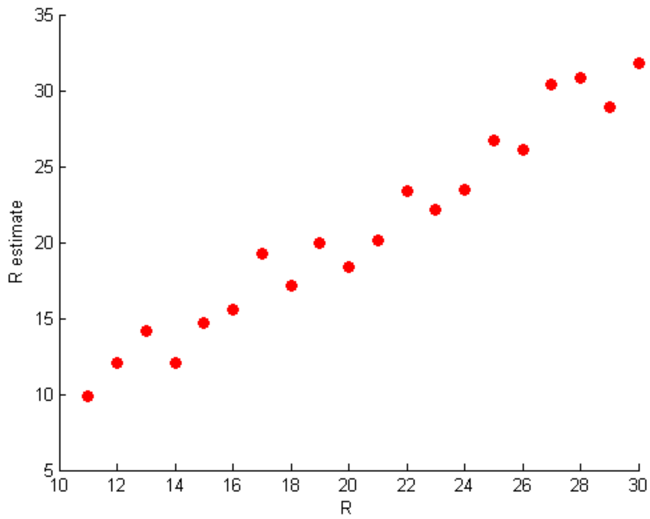


Figure 5.1: Illustration of the performance of the ML estimator for R. The x axis represents the true values whereas the y represents the estimates.

Estimation of α via a Message Passing algorithm

Estimation of α is a stepwise process since it relies on available data , such as distance and connectivity status, but also to another parameter of the model which is R. Another aspect that characterizes the estimation problem of α is time evolution, thus an extra issue is added to the problem. In order to simplify the process of obtaining $\hat{\alpha}$ the problem is separated in different steps. The first step of the followed methodology obtains an estimate of $\hat{R}$ with respect to the process introduced in 5.0.2. As soon as $\hat{R}$ is obtained the next step of the methodology passes the gathered information in order to obtain an estimate upon $\frac{\pi_0}{\pi_1}$ in order to investigate the possible limitations upon the search area for α. The last step of the methodology provides information regarding the distance in between a pair of nodes and the corresponding connectivity status over time. Finally an estimation upon $\hat{\alpha}$ is obtained with respect to the available information and the ML criterion introduced in (5.6). In order to evaluate the performance of the proposed methodology a simulation scenario where two static agents attempt to exchange information is considered. The simulation scenario is repeated for various distances and different values of α and R in order to obtain an overall idea about the different aspects of the problem and the corresponding impact on estimating $\hat{\alpha}$. A visual illustration upon the performance of the proposed methodology is provided in figure 5.3.

The performance of the algorithm varies over he different pairs of (R,α). A basic aspect that affects the accuracy of $\hat{\alpha}$ is the accuracy of $\hat{R}$. In our case for example this holds for the third instance illustrated in figure 5.3. On the other hand A good

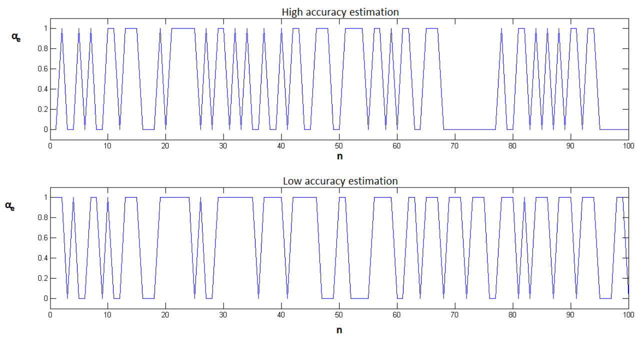


Figure 5.2: Illustration of the most accurate and the poorest performance of ML based on the given data set.

estimate of $\hat{R}$ does not necessarily yield a good estimation for $\hat{\alpha}$ as it can be seen from instance 14 and also a poor estimate of $\hat{R}$ does not necessarily yield a poor estimate for $\hat{\alpha}$ i.e simulation instance 17. Thus a deeper investigation upon the factors that affect the estimation of $\hat{\alpha}$ is required.

The followed methodology relies upon $\hat{R}$ and the main source of information which in that case is the connectivity status over the investigated over the time period. Thus an analysis upon the influence of the available information and the impact upon the estimation of $\hat{\alpha}$ is required. In order to simplify the investigation we focus upon the extreme behaviour of the estimator. In figure 5.4 we illustrate the connectivity status for the cases where the estimator yields the optimal result as well as the case where the estimator introduced the poorest performance.

As it can be seen from the results the two cases follow a very different connectivity pattern. On one hand for the case where the two agents spend most of the time without having a connection, in such case the $1-\beta$ component of the *Markovian* model dominates on the output sequence of states, whereas when the two agents are highly connected over time , in such case $1-\alpha$ dominates the output sequence of states, the algorithm can easier track a good estimate for α.

An interesting aspect though related with the two cases is that the corresponding estimate for $\hat{R}$ is poor in both cases. Nevertheless the estimator yields a good estimate for the second case. Thus obtained information from the available data set can compensate the impact from a poor estimation of $\hat{R}$.

Another issue that can be considered is what happens if we follow the inverse approach and in particular what is the output of the ML criterion for $\hat{\alpha}$ when the estimator of $\hat{R}$ introduces an extreme behaviour. Therefore we investigate the impact of such behaviour along with the corresponding connectivity status. The obtained results are illustrated in figure 5.5. The interesting issue upon this result

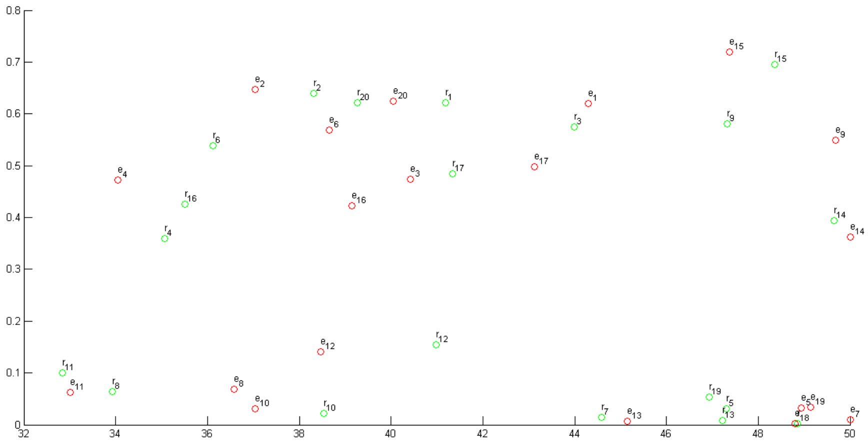


Figure 5.3: An illustration of the performance of the followed methodology for $\hat{R}$. The x axis represents the values for R and $\hat{R}$ whereas the y axis illustrates the region for α and $\hat{a\alpha}$. The original pair of values is represented by the green points denoted as r_i whereas the estimated pairs are illustrated by the red points denoted as e_i for $i = 1 \ldots 20$. A subset of the realizations is provided in order to make the visual comparison feasible.

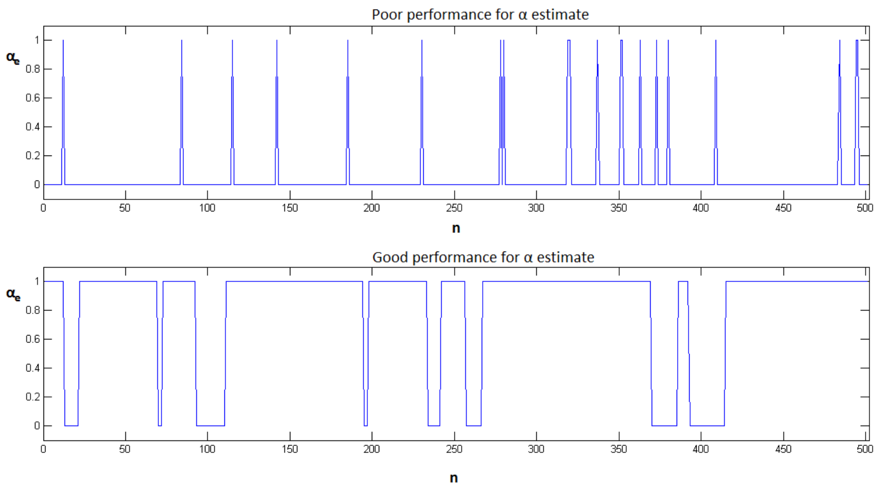


Figure 5.4: Illustration of the connectivity status for the extreme cases of the performance of the estimator. The x axis represents time whereas the y axis represents the connectivity status.

has to do with the impact upon the estimation of $\hat{\alpha}$. The error introduced by the estimator of $al\hat{p}ha$ is slightly higher (0.0190) when the estimator of $\hat{R}$ yields the optimal estimation compared to the error (0.0182) that corresponds to the poorest estimation of $\hat{R}$. The reason why this result occurs is related with the fact that , in case of the optimal performance of $\hat{R}$,the two agents spend most of the time without having a connection with each other thus the component $1 - \beta$ dominates the output state of sequences. On the other hand the connectivity status is dense over the investigated time period thus the provided information compensates the impact from the poor performance from the estimator of $\hat{R}$. Thus, depending on the connectivity status, on one hand the estimator of $\hat{\alpha}$ compensates the impact from a poor estimation of $\hat{R}$ but on the other hand it is likely that the best estimation of $\hat{R}$ given the available does not yield a very good estimation for $\hat{\alpha}$ due to the fact that the available information misleads the corresponding estimator. Derived on this, we can deduce that the estimation of $\hat{\alpha}$ is a combinatorial issue depending on the estimation of the model parameter as well as the available information.

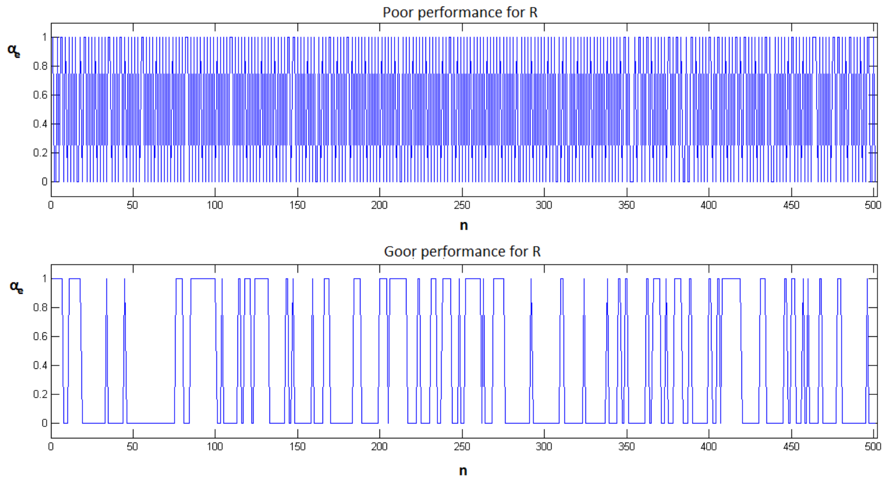


Figure 5.5: Representation of the connectivity status in cases where the estimator of $\hat{R}$ introduces an extreme performance. The upper part of the figure corresponds to the poorest estimation whereas the lower part corresponds to the optimal estimation.

The current section investigated the performance of the estimators considering the problem of the model parameters as an estimation problem. The main interest upon this investigation focus on the different aspects of the estimation problem such as the available data set as well as the performance of a preceding estimator in case of the α parameter. The simulation scenario focus on single pair of nodes so that the size of the estimation problem provides us the opportunity to conduct an in depth analysis easier. A basic requirement when conducting such research is that different data sets are available for each of the parameters. Such a scenario is rare in real life campaign measurements thus the investigation on methods to estimate the model parameters continues with methods where the parameters can be estimated jointly by the same data set.

5.1 Joint estimation of R,α single link static nodes

The so far approach to acquire the model parameters introduced a message passing algorithm in order to estimate the pair of parameters of R,α. At each step of the algorithm the ML criterion is applied for the corresponding parameter. The obtained results indicated that the accuracy of the output relies on several factors such as the measurements or , in case of α, on the accuracy of the estimator that preceded it. Another outcome of the followed approach is that even in cases where the output of the ML criterion for $\hat{R}$ results a poor estimation the corresponding result for $\hat{\alpha}$ can be compensated depending on the available data set, hence even in cases where the estimator of R yields a poor performance the estimator of $alpha$ may yield a more accurate estimate due to the available data. Thus the corresponding estimator performs better when more factors are taken into account. Derived on this it is expected that an estimator where the pair of parameters is jointly obtained yields a more accurate estimation. Given a set of data $A(n)$ where:

$$A(n) = (a_e(n)|D(n)). \tag{5.14}$$

The joint distribution for a single link reads:

$$f(A, \alpha, R) = f(A|\alpha, R)f(\alpha, R) = f(A|\alpha, R) \cdot f(\alpha|R) \cdot f(R). \tag{5.15}$$

where each of the introduced factors takes the following form:

1. $f(R) \sim U(o, p)$.
2. $f(\alpha|R) \sim U(0, \alpha_{max})$.
3. $f(A|\alpha, R) = s_0 \cdot \alpha^{|A|} \cdot (1-\alpha)^{|\Delta|} \cdot \prod_{n \in B} (\alpha \cdot \frac{e^{\frac{-D(n)^2}{2 \cdot \hat{R}^2}}}{1-e^{\frac{-D(n)^2}{2 \cdot \hat{R}^2}}}) \cdot \prod_{n \in \Gamma} (1 - \alpha \cdot \frac{e^{\frac{-D(n)^2}{2 \cdot \hat{R}^2}}}{1-e^{\frac{-D(n)^2}{2 \cdot \hat{R}^2}}})$.

A widely used approach to represent statistical relationships graphically are factor graphs. The factor graph, for a single link, that represents the joint distribution introduced in equation (5.15) is illustrated in figure 5.6.

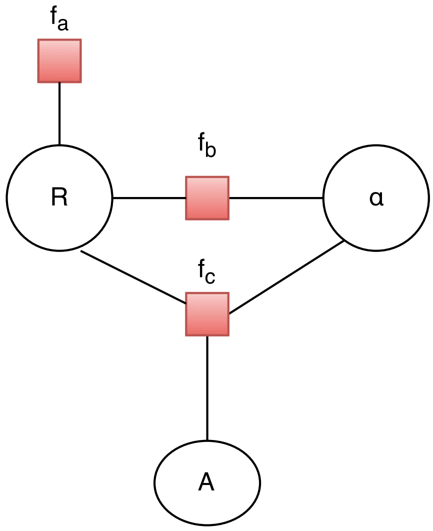


Figure 5.6: Graphical representation of the joint distribution $f(A, \alpha, R)$. The corresponding factors read: $f_a = f(R)$, $f_b = f(a|R)$,$f_c = f(A|\alpha, R)$.

The estimate upon the pair of parameters R, α is obtained by maximizing the output of MAP estimator introduced in equation (5.15). Derived on this, the estimator takes the following form:

$$(\hat{R}, \hat{\alpha}) = \arg\max_{R,\alpha} f(A|\alpha, R) \cdot f(\alpha|R) \cdot f(R). \tag{5.16}$$

Based on the fact that the current form of the quantity R does not provide any particular information about the corresponding parameter whereas $\alpha|R$ provides information only about the size of the corresponding search region, the rule introduced in equation (5.16) neglects to the ML criterion introduced in the following equation:

$$(\hat{R}, \hat{\alpha}) = \arg\max_{R,\alpha} f(A|\alpha, R). \tag{5.17}$$

The joint estimate for R, α is obtained when the following criterion is fulfilled:

$$\nabla f(A|\alpha, R) = J = \left[\frac{\partial f(A|\alpha, R)}{\partial R}, \frac{\partial f(A|\alpha, R)}{\partial \alpha}\right]^T = 0. \tag{5.18}$$

By taking into account the form of $f(A|\alpha, R)$ the equation introduced in (5.18) for the static case ,where the index n from $D(n)$, the first component of the gradient vector J can be rewritten as follows:

$$J_{11} = \frac{D^2}{R^3} \cdot \alpha^{|A|} \cdot (1-\alpha)^{\Delta} \Bigg[(-1)^{\nu} \cdot e^{\frac{-D^2}{2 \cdot R^2}} \cdot \left(\alpha \cdot \frac{e^{\frac{-D^2}{2 \cdot R^2}}}{1 - e^{\frac{-D^2}{2 \cdot R^2}}} \right)^{|B|} \cdot \left(1 - \alpha \cdot \frac{e^{\frac{-D^2}{2 \cdot R^2}}}{1 - e^{\frac{-D^2}{2 \cdot R^2}}} \right)^{|\Gamma|}$$
$$+ |B| \cdot s_0 \cdot \alpha \cdot \left(\alpha \cdot \frac{e^{\frac{-D^2}{2 \cdot R^2}}}{1 - e^{\frac{-D^2}{2 \cdot R^2}}} \right)^{|B|-1} \cdot \frac{e^{\frac{-D^2}{2 \cdot R^2}}}{(1 - e^{\frac{-D^2}{2 \cdot R^2}})^2} \cdot \left(1 - \alpha \cdot \frac{e^{\frac{-D^2}{2 \cdot R^2}}}{1 - e^{\frac{-D^2}{2 \cdot R^2}}} \right)^{|\Gamma|}$$
$$- s_0 \cdot \alpha \cdot \left(\alpha \cdot \frac{e^{\frac{-D^2}{2 \cdot R^2}}}{1 - e^{\frac{-D^2}{2 \cdot R^2}}} \right)^{|B|} \cdot \frac{e^{\frac{-D^2}{2 \cdot R^2}}}{(1 - e^{\frac{-D^2}{2 \cdot R^2}})^2} \cdot |\Gamma| \cdot \left(1 - \alpha \cdot \frac{e^{\frac{-D^2}{2 \cdot R^2}}}{1 - e^{\frac{-D^2}{2 \cdot R^2}}} \right)^{|\Gamma|-1} \Bigg] = 0. \tag{5.19}$$

where ν is defined as follows:

$$\nu = \begin{cases} 0, \text{if } s_0 = \pi_0 \\ 1, \text{if } s_0 = \pi_1 \end{cases} \tag{5.20}$$

whereas the second component of the gradient vector J is written as follows:

$$\begin{aligned} J_{12} = s_0 \cdot (&|A| \alpha^{|A|-1} \cdot (\alpha \cdot \frac{\pi_0}{\pi_1})^{|B|} \cdot (1 - \frac{\pi_0}{\pi_1} \cdot \alpha)^{|\Gamma|} \cdot (1-\alpha)^{|\Delta|} \\ &+ \alpha^{|A|} \cdot \frac{\pi_0}{\pi_1} \cdot |B| \cdot (\alpha \cdot \frac{\pi_0}{\pi_1})^{|B|-1} \cdot (1 - \frac{\pi_0}{\pi_1} \cdot \alpha)^{|\Gamma|} \cdot (1-\alpha)^{|\Delta|} \\ &- \alpha^{|A|} \cdot (\alpha \cdot \frac{\pi_0}{\pi_1})^{|B|} \cdot |\Gamma| \frac{\pi_0}{\pi_1} \cdot (1 - \frac{\pi_0}{\pi_1} \cdot \alpha)^{|\Gamma|-1} \cdot (1-\alpha)^{|\Delta|} \\ &- \alpha^{|A|} \cdot (\alpha \cdot \frac{\pi_0}{\pi_1})^{|B|} \cdot (1 - \frac{\pi_0}{\pi_1} \cdot \alpha)^{|\Gamma|} \cdot |\Delta| \cdot (1-\alpha)^{|\Delta|-1}) = 0. \end{aligned} \tag{5.21}$$

In order to obtain an intuition regarding the ML criterion for the joint distribution, a simple experiment is conducted. The trivial static nodes connectivity scenario over a time period of 600 time instances. The corresponding results are illustrated in figure 5.7.

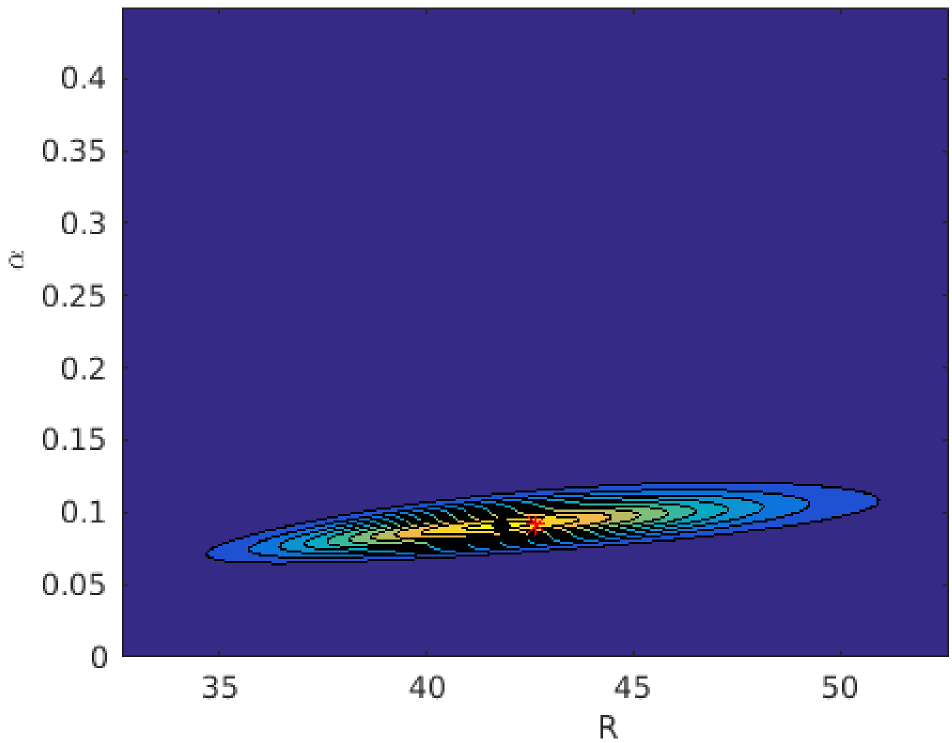


Figure 5.7: Illustration of the results obtained by the joint estimation of the parameters $\hat{R}, \hat{\alpha}$. The true values of the parameters are represented by the red cross whereas the pair of estimates is illustrated by the black circle.Search region=[35,55], $\alpha = 0.9$,$R = 43$

5.1.1 Joint estimation of R, α multiple links-dynamic case

So far we introduced the joint estimation for $\hat{\alpha}, \hat{R}$ via ML for a single communication link. The pairwise connectivity in between different nodes is considered to be independent from on pair to the other. Consequently the extension of the ML function for a single communication link to the Likelihood functions for a graph is straight forward since the total ML function is the product from all the pairwise Likelihood functions.The Likelihood function for the whole graph can be expressed as follows:

$$\Lambda(\alpha, R|A) = \prod_{i>j} \prod_{n=0}^{N} P(a_{e_{ij}}(n) = a_{e_{ij}}^{(d)}(n)|(x_i(n), x_j(n))^{(d)}, \alpha, R) \tag{5.22}$$

With respect to the general formulation of Likelihood function, the analytical form of the function can then be expressed as follows:

$$\Lambda(\alpha, R|A) = \alpha^{|A|}(1-\alpha)^{|\Delta|} \prod_{a_{e_{ij}}(0)=1} e^{\frac{-Dij^2}{2R^2}} \prod_{a_{e_{ij}}(0)=0} (1 - e^{\frac{-Dij^2}{2R^2}})$$
$$\prod_{n\in\Gamma}(1 - \alpha \cdot \frac{e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}{1 - e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}) \cdot \prod_{n\in B} \alpha \cdot \frac{e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}{1 - e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}} \tag{5.23}$$

The estimate upon the parameters of R, α corresponds to the pair of of values that maximize the likelihood function:

$$(\hat{\alpha}, \hat{R}) = \arg\max_{R,\alpha} \Lambda(\alpha, R) \tag{5.24}$$

Alternatively the estimate upon the pair of parameters can be obtained by maximizing the logarithm of ML. In such case the corresponding function reads:

$$ln\Lambda(\alpha, R) = \sum_{i>j}\Big[|A|ln\alpha + |\Delta|ln(1-\alpha) + \sum_{a_{e_{ij}}(0)=1} \frac{-Dij(n)^2}{2R^2} + \sum_{a_{e_{ij}}(0)=0} ln(1 - e^{\frac{-Dij(n)^2}{2R^2}})$$
$$+ \sum_{n\in\Gamma} ln(1 - \alpha \cdot \frac{e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}{1 - e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}) + |B|ln\alpha + \sum_{n\in B} \frac{-Dij(n)^2}{2R^2} - \sum_{n\in B} ln(1 - e^{\frac{-Dij(n)^2}{2R^2}})\Big] \tag{5.25}$$

With respect to the log likelihood the *Jacobian* vector is then written as follows:

$$J = \Big[\frac{dln\Lambda(\alpha, R)}{dR}, \frac{dln\Lambda(\alpha, R)}{d\alpha}\Big]^T \tag{5.26}$$

where J_{11} reads:

$$\frac{dln\Lambda(\alpha, R)}{dR} = \sum_{i>j}\Big[\sum_{a_{e_{ij}}(0)=1} D_{ij}(0)^2 - \sum_{a_{e_{ij}}(0)=0} \frac{e^{\frac{-D_{ij}(0)^2}{2\cdot R^2}}}{1 - e^{\frac{-D_{ij}(0)^2}{2\cdot R^2}}} + \sum_{n\in B} D_{ij}(n)^2$$
$$+ \sum_{n\in B} \frac{e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}{1 - e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}} - \alpha \sum_{n\in\Gamma} \cdot \frac{D_{ij}^2 \cdot e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}{(1 - (\alpha+1)\cdot e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}})(1 - e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}})}\Big] \tag{5.27}$$

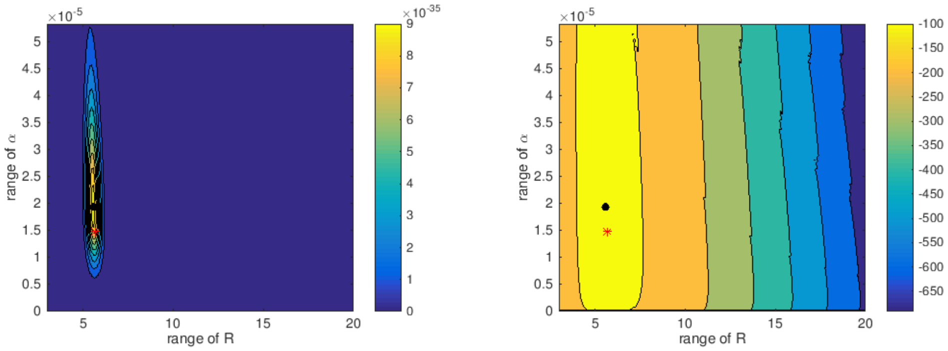
and J_{21} reads:

$$\frac{dln\Lambda(\alpha, R)}{d\alpha} = \sum_{i>j}\Big[\frac{|A|}{\alpha} + \frac{|B|}{\alpha} - \frac{|\Delta|}{1-\alpha} - \sum_{n=0}^{N} \frac{e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}{1 - e^{\frac{-D_{ij}(n)^2}{2\cdot R^2}}}\Big] \tag{5.28}$$

Table 5.2: My caption

True value	Searching Grid	Mean estimate	bias	Variance
R_t	Search region	$\mathbb{E}(\hat{R}_{ML})$	$\mathbb{E}(\hat{R}_{ML}) - R$	$\mathbb{E}[\mathbb{E}(\hat{R}_{ML}) - \hat{R}_{ML})^2]$
4.57km	(1,10)km	3.87km	-0.7km	2.96 km^2

In order to obtain a visual inspection upon the behaviour of the ML function over the graph, a simple experiment is conducted. A pair of parameters is arbitrarily picked and the ML function estimates the corresponding values based on the available data set (in such case positions and the corresponding connectivity status over time). The obtained results are illustrated in figure 5.8.



(a) Illustration of the performance of the ML function over the graph

(b) Illustration of the output of the log likelihood function over the graph

Figure 5.8: The current figure illustrates the output of the ML and *log* ML function over the graph. The pair of true values is illustrated by the red cross.

5.1.2 Overview of the performance of the ML estimator via Monte Carlo simulations

There exist several approaches to estimate the parameters of the *Markovian* model. The current work proposed an ML estimator upon the corresponding parameters to address this issue. Several examples have been illustrated so that we can have an idea about the performance of the estimator. In order to obtain a comprehensive idea about the performance of the ML estimator though, an extended experiment is required. Therefore a number of 100 Monte Carlo simulations is conducted. A pair of true values (R_t, α_t) is arbitrarily picked and at each realization the estimator attempts to provide as accurate estimates as possible by taking into account different data sets. The obtained results as well as the rest of the elements considered in each realization are illustrated individually for each of the parameters in the following tables:

Table 5.3: My caption

True value	Mean estimate	bias	Variance
α_t	$\mathbb{E}(\hat{\alpha}_{ML})$	$(\mathbb{E}(\hat{\alpha}_{ML}) - \alpha_t)$	$\mathbb{E}[(\mathbb{E}(\hat{\alpha}_{ML}) - \hat{\alpha})^2]$
$1.56{\cdot}10^{-4}$	$3.78{\cdot}10^{-4}$	$2.2{\cdot}10^{-4}$	$1.37{\cdot}10^{-7}$

The results indicate the bias of the estimates is relatively high. In order to obtain a better understanding why this phenomenon occurs a visual inspection over the different pair of estimates obtained by the Monte Carlo simulations is provided in figure 5.9.

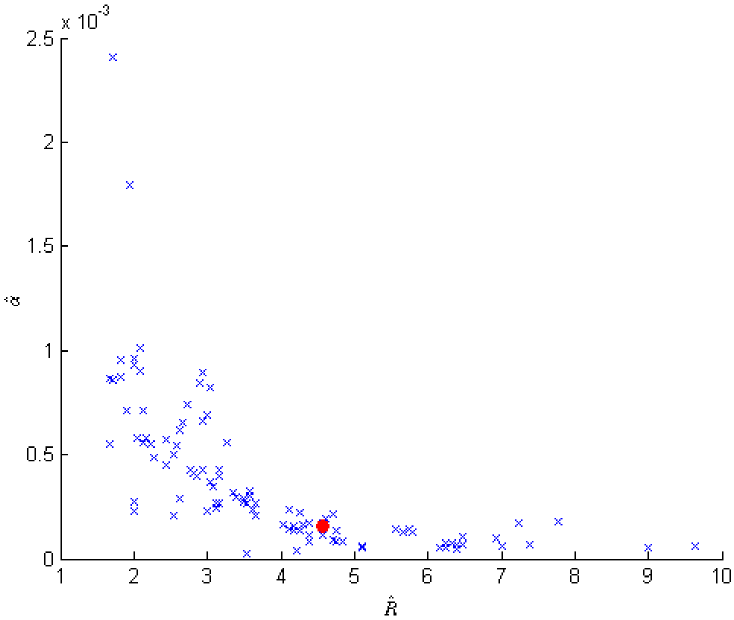
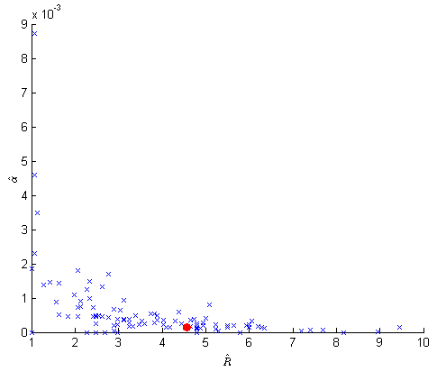
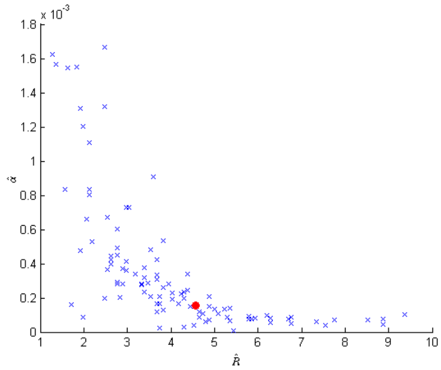


Figure 5.9: Illustration of the results obtained by the joint estimation of the parameters $\hat{R}, \hat{\alpha}$ via the ML estimator. The red point corresponds to the true pair of parameters of the *Markovian* model. .

The illustrated results indicate that the performance of the estimator is affected by the dense presence of outliers. The reason why this phenomenon occurs arises from the fact that the estimator needs a larger amount of data to eliminate the affect from the outliers and make more accurate estimates. In order to obtain a better understanding about the performance of the estimator over the different size of data sets the same simulation set up is repeated for a fewer time instances. The initial experiment took investigated the evolution of the connectivity status over 300 time instances. The size of the data set shortens to 50 and 200 time instances accordingly. The obtained results are illustrated in figure 5.10.



(a) Estimates over 50 time instances



(b) Estimates over 200 time instances

Figure 5.10: The current figure illustrates how the estimates are spread over the search grid based on different sizes of data set.

The results indicate while the size of the data shortens the presence of outliers becomes more dense. The particular outcome is expected since the available information becomes less while the time period shortens. The current work considers a time period of 300 time instances ao that the simulation task is computationally feasible in an amount of time that serves the purpose of the current work.

The same simulation pattern introduced before is followed but now instead of changing the size of the data, 300, we change the R parameter over a wide grid. In that way we obtain a more comprehensive idea regarding the performance of the ML estimator considering possible applications on other systems. Another aspect when considering R has to do with the fact that it affects the sparsity on the graph. Therefore we are seeking to affect the affect on the performance of the ML criterion under different connectivity conditions. Hence, we conduct the same simulation pattern over a wide grid for R over $[1, 36]$. It is expected that the graph will become less sparse while R increases. The obtained results are illustrated in figure 5.11.

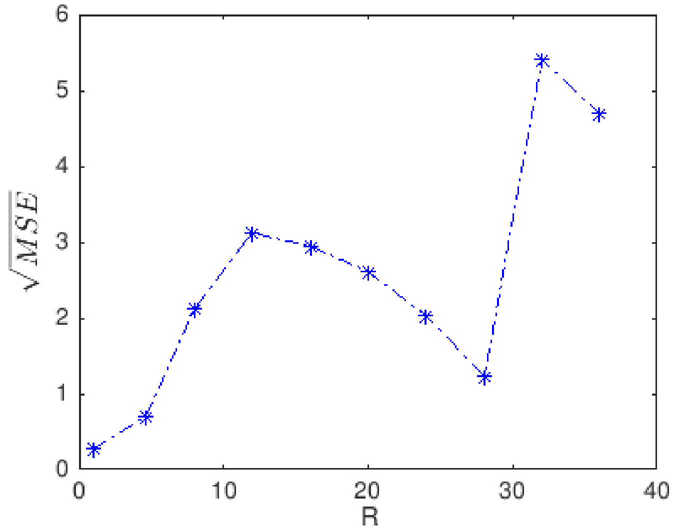


Figure 5.11: Illustration of the results obtained by the simulations over the grid [1,36].

.

The obtained results indicate that the error illustrates a variation over the grid, i.e it does not yield a linearly increasing tendency. Considering the fact that higher R corresponds to a higher level of connectivity the obtained results indicate that there is some kind of trade off in between the level of connectivity and the error. The error illustrates a linearly increasing tendency for a relatively low level of connectivity even though R increases while the opposite outcome holds when the level of connectivity is midrange. The error indicates a tendency to increase very fast for high levels of connectivity. The size of the error per point of R could be addressed as an indicator regarding the size of the required data so that the estimator can make more accurate estimates.

5.2 Model implementation

A well known technology used for data transmission is the Global System for Mobile communications (GSM). GSM is a standard developed to describe the protocols for digital cellular networks used for mobile phones. The transmission range for cellular networks is from 1 km to 20 km. In cases of urban areas signal propagation is severely affected by the surroundings, consequently the usage of cell towers with a range that reaches the upper bound becomes impractical in dense areas. Therefore a common practise is to place several base stations, closely spaced with each other, with short or midrange transmission capacity. In such case the transmission range varies from 1 km to 5 km. Cooperative Localization paradigm aims to overcome the impracticalities of the current technologies which mainly arise in dense areas. Thus the focus of the current work is on this cases along with all the corresponding information available, which in that case is the transmission range of the areas of interest.

Considering all these aspects from the perspective of the connectivity model, the corresponding information can be utilised in order to develop a mathematical framework that captures the model parameter related to distance which in such case is R. One approach to model phenomena with the particular type of properties is the *Rayleigh* distribution. There exist other possible candidates, i.e the *gamma* distribution, which can also represent these properties. In our case though the available information is limited thus it is desired to develop a model a simple as possible. Thus the usage of the *Rayleigh* distribution is preferable since it is a single parameter model. Derived on this the R parameter can then be represented as follows:

$$R \sim Rayleigh(\sigma), \text{where } \sigma \text{ is the scale parameter of the distribution.} \tag{5.29}$$

The probability density function of R takes then the following form:

$$f(x;\sigma) = \frac{x}{\sigma^2} \cdot e^{\frac{-x^2}{2\cdot\sigma^2}},\ x \geq 0. \tag{5.30}$$

The σ parameter is specified with respect to the available information about the system. The choice of σ reflects on the form of the distribution. In cases where a very good knowledge about the system is available, this information can be expressed by a distribution with a scale parameter that introduces a narrow pick in the corresponding pdf. Consequently the use of an estimator that utilizes the prior information about the system, such as a Maximum a Posteriori estimator, is affected by the prior pdf. On the other hand in cases where the information about the system is not in detail or it is spread over the search region, it is preferable to formulate the distribution with a scale parameter that spreads the probability mass over the search region. Based on these characteristics for σ it is important to make a setting for the scale parameter that leads to representative formulation of the pdf.

In order to obtain a visual inspection about the form of the distribution of R with respect to the proposed model an illustration of the model is provided in figure 5.12.

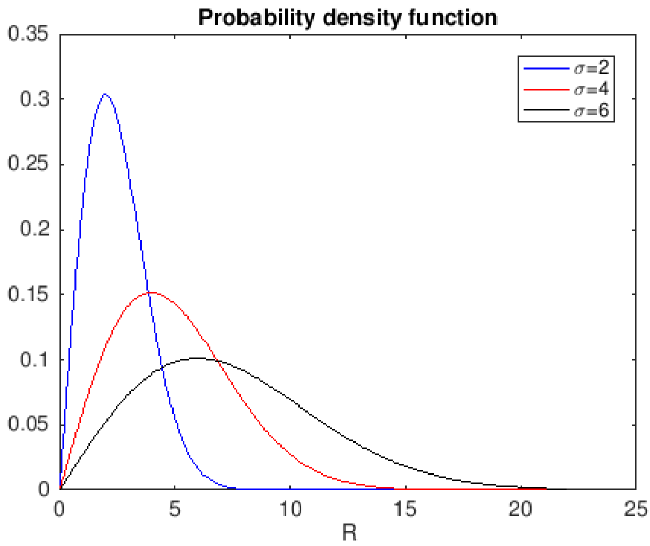


Figure 5.12: Prior probability for R for different settings of σ.

5.2.1 Prior information about the α parameter

The parameter that forms the original model as a stationary distribution, in such case α, expresses how likely a pairwise connection is about to drop. Considering the parameter in a real life environment, α is the parameter that reflects the effect of the area of interest in the signal propagation. In cases where the area of interest consists of tall buildings, tunnels or a subway then it is quite likely that the users will loose connection quite often. On the other hand, in cases where the area of interest is a plain terrain or a park then it is expected that the phenomenon of loosing connection is rare. Derived on this setting the α parameter requires good knowledge about the environment since there can be a significant different upon the selection of α from the one case to the other.

A very commonly used prior for a probability variable is the *Beta* distribution. Since the α parameter expresses such a variable modelling the prior info about the parameter as a *Beta* distribution is an obvious choice:

$$\alpha \sim Beta(a,b), \text{where } a,b \text{ are the shape parameters of the distribution.} \tag{5.31}$$

Where the probability density function of α takes then the following form:

$$f(x;a,b) = \frac{x^{(a-1)} \cdot (1-x)^{(b-1)}}{B(a,b)}. \tag{5.32}$$

where $B(a,b)$ the so called *Beta* function expressed as:

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt. \tag{5.33}$$

The fact that the model consists of two parameters indicates that a very good prior knowledge about α is required. A special case of the Beta distribution with $a = b = 1$ is the Uniform distribution. In order to obtain a visual inspection about the possible form of the pdf of α considering the *Markovian* model, three characteristic cases of the model are illustrated in figure 5.13.

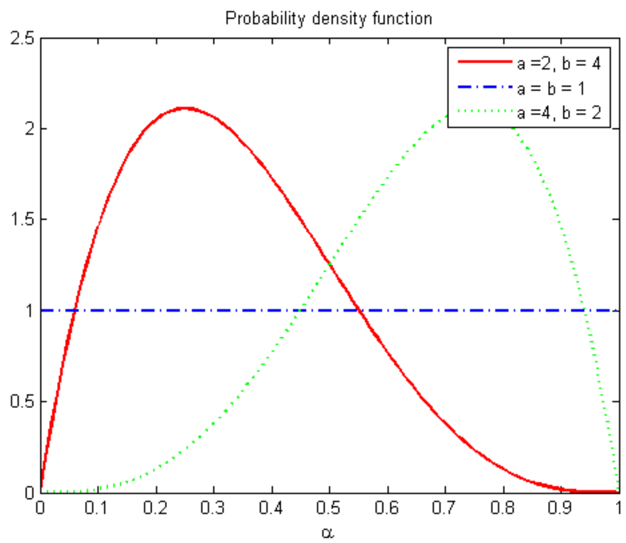


Figure 5.13: The current figure illustrates three characteristic cases for the pdf of α. The red curve corresponds to the case where is less likely to fail whereas the green curve corresponds to the opposite scenario. On the other hand the blue curve corresponds to the exceptional case of the uniform distribution when in practise no prior information is available.

The current work does not have any particular knowledge regarding a specific case for alpha thus in for our study we are considering that α follows a uniform distribution. In such case α is represented as follows:

$$\alpha \sim Beta(1,1). \tag{5.34}$$

However given further related system information leads to a different setting of the parameters of the prior pdf for α.

5.2.2 Estimation of $(\hat{R}, \hat{\alpha})$ via a MAP estimator

The current work considers the GSM technology for data transmission. Derived on this fact one extra piece of information upon the joint estimation over the graph. In equation 5.16 the joint distribution for a single pair of nodes is introduced where the prior information about is considered to follow a uniform distribution. Thus an ML estimator is developed over the graph that is represented in equation (5.22). The consideration upon the GSM technology and the mapping of the corresponding knowledge upon urban areas to a *Rayleigh* model adds a prior information about R that can be taken into account. Consequently the ML estimator introduced in equation can be extended to a Maximum a Posteriori estimator by inserting the corresponding piece of information to the function. To do so the corresponding prior pdf is simply inserted to the corresponding factorization. Unlike R, the modelling of α as a uniform distribution does not contribute to any additional knowledge thus the corresponding pdf is neglected. The general formulation for the Maximum a Posteriori Probability reads:

$$P(\theta|x) = \frac{P(x|\theta)}{P(x)} = \frac{P(x|\theta) \cdot P(\theta)}{P(x)} \propto P(x|\theta) \cdot P(\theta). \tag{5.35}$$

Based on these considerations and with respect to the equation introduced in equation (5.30) the MAP estimator for the *Markovian* model is expressed as follows:

$$(\hat{\alpha}, \hat{R})_{MAP} = \arg\max_{\alpha,R} \mathbb{1}(\alpha \in [0,1]) \cdot \frac{R}{\sigma^2} \cdot e^{\frac{-R^2}{2\cdot\sigma^2}}$$
$$\cdot \prod_{i>j} \prod_{n=0}^{N} P(a_{e_{ij}}(n) = a_{e_{ij}}^{(d)}(n) | (x_i(n), x_j(n))^{(d)}, \alpha, R)$$

Comparison between the ML estimator and the MAP estimator

The consideration upon the GSM technology leads eventually to the development of a the MAP estimator which adds one extra tool to the estimation of the model parameters. As soon as the form of the MAP estimator is defined the next step in our work is to evaluate what are the profits of having a prior knowledge about the system compared to the ML approach that has been followed so far. Therefore the experiment which is conducted to obtain an overall idea about the performance of ML is also conducted for the MAP estimator. Consequently the two estimators are compared upon identical criteria. Note for the prior pdf of R, $\sigma = 2.4$ and $\alpha \sim Beta(1,1)$. A visual inspection upon the behaviour of the corresponding pdf can be obtained in figure 5.12. The results of the performance of the MAP estimator are represented in tables 5.4 and 5.5 accordingly.

Table 5.4: Results from the MAP estimator for the R parameter

True value	Searching Grid	Mean estimate	bias	Variance
R_t	Search Region	$\mathbb{E}(\hat{R}_{MAP})$	$\mathbb{E}(\hat{R}_{MAP}) - R_t$	$\mathbb{E}[\mathbb{E}(\hat{R}_{MAP}) - \hat{R}^2_{MAP}]$
4.57km	(1,10)km	3.84km	-0.73km	2.82km^2

Table 5.5: Results from the MAP estimator for the α parameter

True value	Mean estimate	bias	Variance
α_t	$\mathbb{E}(\hat{\alpha}_{MAP})$	$\mathbb{E}(\hat{\alpha}_{MAP}) - \alpha$	$\mathbb{E}[(\mathbb{E}(\hat{\alpha}_{MAP}) - \hat{\alpha})^2]$
$1.6 \cdot 10^{-4}$	$3.8 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$	$1.36 \cdot 10^{-7}$

A visual inspection upon the obtained estimates over the different Monte Carlo simulations is illustrated in figure 5.14.

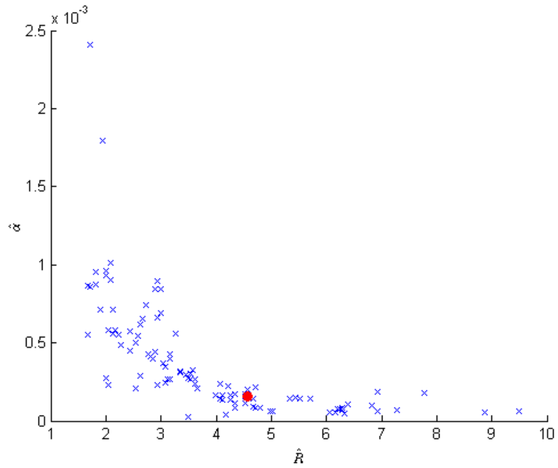


Figure 5.14: Illustration of the results obtained by the joint estimation of the parameters $\hat{R}, \hat{\alpha}$ via the MAP estimator. The red point corresponds to the true pair of parameters of the *Markovian* model. .

The results indicate that no particular benefit can be obtained from the MAP estimator due to the prior information upon R. That kind of result is expected since the available data set from the ML function dominates the factorization and as a result the contribution from the prior pdf has a minor influence to the corresponding output.

Chapter 6

Model evaluation and performance

The so far conducted research of the current work investigates the connectivity of a dynamic network. An analysis upon the advantages and the disadvantages of each of the different approaches have been discussed through but a straight forward comparison is not feasible on a particular domain since it is difficult to compare the approaches i.e topologically. The same problem arises when attempting to evaluate the performance of the ML estimator since no default approach exists to address the estimation issue over a dynamically involved network. One way to address both cases is by make the comparison on the time domain.

A main issue that arises when attempting to conduct such a comparison though is the type of methodology that can be followed to investigate the time evolution of a graph. The corresponding literature is characterised by the lack of relative methodologies. One approach that can be followed to address that kind of issue is the so called *Jaccard* index [9] also known as *Jaccard* coefficient which is a statistic utilized to compare the similarity and diversity of sample sets. The main idea that relies upon the *Jaccard* coefficient is that the correlation in between the different sets is expressed as a ratio in between the common elements of the sets over the total number of elements included in both sets. The corresponding relation considering that the sets investigate the similarities over the edge sets $\mathcal{E}_1$, $\mathcal{E}_2$ can then be expressed as follows:

$$J(\mathcal{E}_1, \mathcal{E}_2) = \frac{|\mathcal{E}_1 \cap \mathcal{E}_2|}{|\mathcal{E}_1 \cup \mathcal{E}_2|}, \text{where } 0 \leqslant J(\mathcal{E}_1, \mathcal{E}_2) \leq 1 \tag{6.1}$$

There exist a special case where both sample sets are empty. In such case we define $J(\mathcal{E}_1, \mathcal{E}_2) = 1$.

An alternative approach that relies upon the same idea but instead of the similarities it measures the dissimilarity in between sample sets is the so called *Jaccard* distance which is expressed as follows:

$$d_J = 1 - J(\mathcal{E}_1, \mathcal{E}_2) = \frac{|\mathcal{E}_1 \cup \mathcal{E}_2| - |\mathcal{E}_1 \cap \mathcal{E}_2|}{|\mathcal{E}_1 \cup \mathcal{E}_2|} \tag{6.2}$$

6.1 Graph Consistency based on the *Jaccard* index

Once the correlation methodology is specified the next step is to investigate the time correlation in between the graphs of particular set up and conduct an analysis over the different states. The current simulation scenario is the trivial midrange simulation scenario. We are mainly interested for the case where the case where a communication link is established in between two agents and how this relationship evolves over time. Therefore when we consider the *Jaccard* metric over k different time instances and for the current simulation scenario we investigate the correlation in between the first time instance and the $k-1$ time instances that follow. Thus the *Jaccard* component for the current simulation scenario can be expressed as follows:

$$J(\mathcal{E}_1, \mathcal{E}_{1+j}) = \frac{|\mathcal{E}_1 \cap \mathcal{E}_2|}{|\mathcal{E}_1 \cup \mathcal{E}_{1+j}|}, \text{for } j = 1, \ldots k. \tag{6.3}$$

For the current scenario we investigate the consistency of the graph for the *Markovian* connectivity model as well as the connectivity model of *Henk Wymeersch*. The parameters for each of the connectivity approaches as well as the components for the particular set up are as follows:

Table 6.1: Simulation parameters

realizations	time instances	α	Radius	R
100	100	0.7	20	13

The obtained results are illustrated in figure 6.1.

The results indicate that *Jaccard* metric captures the relationship in between two consecutive time instances pretty well. On one hand the *Jaccard* metric illustrates the relationship in between graphs but on the other hand we cannot have an idea how the graphs differ over time. Considering the case where for example the number of participants in the graph remains the same over time, the metric yields the number of connection that remain the same over time but there is no information about the connectivity status regarding the cases where the connectivity status was different from one case to the other. In order to provide a better intuition regarding this consideration two trivial examples are illustrated in figures 6.2 6.3 where we illustrate how two different networks evolve over time and what is the connection in between them.

The first example considers the scenario where a possible network with three agents transits from a relatively sparse state to a state where all the agents are fully connected with each other.

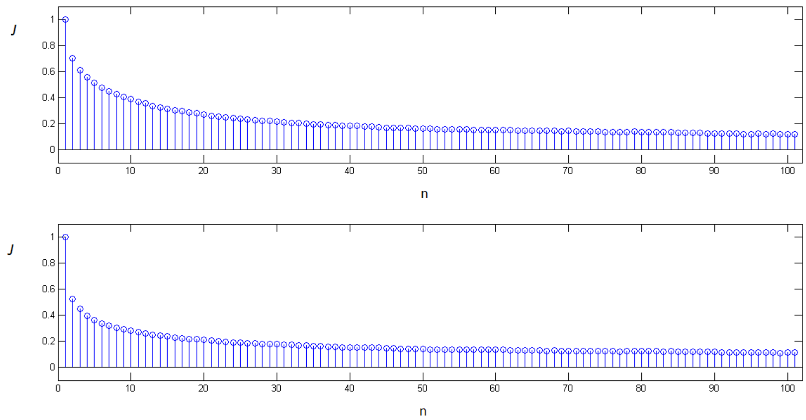
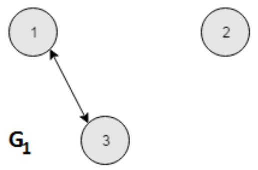
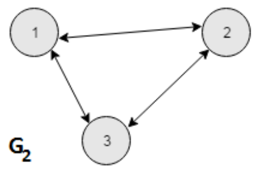


Figure 6.1: Illustration of the *Jaccard* index for the connectivity models. The y axis corresponds to the *Jaccard* component whereas the x represents the time instances. The first subfigure illustrates the results for the connectivity model of *Henk Wymeersc* while the second subfigure illustrates the results for the *Markovian* model.



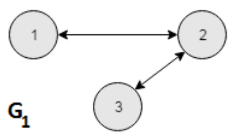
(a) The location of the base nodes along with the transmission radius



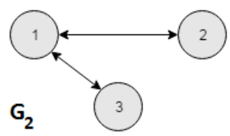
(b) The area where the mobile agent is connected with the base nodes simultaneously.

Figure 6.2: An illustration of the case where from a state where a single bidirectional communication link the first time instance evolves to full network in between the agents.$J(\mathcal{E}_1, \mathcal{E}_2) = \frac{1}{3}$.

Whereas the second example considers the case where the connectivity status in between the agents alters from one state to the other but the total number of connection remains the same.



(a) First time instance of the graph



(b) Second time instance of the graph

Figure 6.3: An illustration of an example where a communication link drops in the meantime another link is generated consequently the total number of connections remains unaltered.$J(\mathcal{E}_1, \mathcal{E}_2) = \frac{1}{3}$.

In both cases the *Jaccard* index yields the same output even though the two graphs evolved in a different manner. Thus there is a need for an additional metric that provides information regarding the rate that the sparsity on the graph evolves over time.

6.2 Jaccard index for the probabilistic model of *Savic* and *Zazo*

The current section investigates the evolution of the connectivity on a dynamic network with respect to the model of *Savic* and *Zazo*. The correlation in between the graphs is estimated with respect to the *Jaccard* index. We follow the midrange connectivity simulation scenario to conduct the analysis. A number of 100 Monte Carlo simulations is conducted in order to obtain a comprehensive idea about the performance of the model. For the current simulation scenario the range parameter is arbitrarily chosen. The obtained results are illustrated in figure 6.4

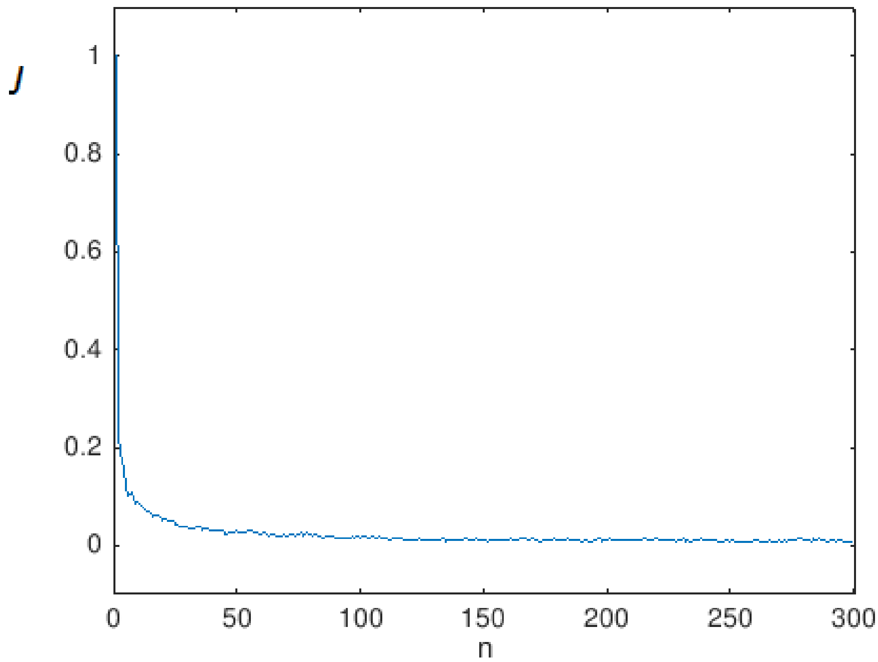


Figure 6.4: Illustration of the performance of the model of *Savic* and *Savic* on a dynamic network.

The results indicate that there is a dependency in between the first time instance and the one that follows. Furthermore the results illustrate that the output of the model yields a relationship which becomes stationary over time. Considering the main idea upon the model is based, meaning that the connection is independent on time, thus the model should be structured in a way that the time aspect is considered.

The *Markovian* model states the idea that the output of the initial model forms a stationary distribution of a *Markov* chain. Based on this approach each instance of the model depends on the instance that preceded it. The particular approach meets the obtained results from the original approach. Derived on this we can claim that the proposed model generalizes the original model on the time domain and overcomes the impracticalities introduced by the initial approach.

6.3 Comparison between the connectivity model of *Henk Wymeersch* and the *Makrovian* model

The connectivity model of *Henk Wymeersch* introduces a straight forward criterion upon we can deduce whether a communication link can be established or not. Even though it is characterised by the obvious drawback that the impact of the surroundings is not taken into account there are several cases where the connectivity model of *Henk Wymeersch* introduces a sufficient approach to address the connectivity issue. In cases where for example the user is very close to the source of the signal or in plane areas where there are not many obstacles in the propagation environment. Thus we seek to investigate whether the *Markovian* model can adopt the behaviour of the connectivity model of *Henk Wymeersch*. The comparison cannot take place topologically therefore we use the *Jaccard* index in order to compare the two approaches on the time domain. The preliminary analysis upon the behaviour of the connectivity model of *Henk Wymeersch* over time indicated that the corresponding distribution tends to become stationary at the first 100 instances. An illustration of this analysis is provided in figure 6.5.

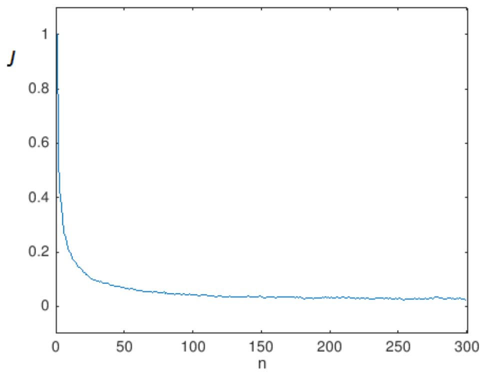


Figure 6.5: Illustration of the behaviour of the connectivity model of *Henk Wymeersch* over time. Note that for the current example $R_H = 10$.

One of the advantages of the *Markovian* model is the flexibility due to the dual parametric structure. For each case upon the performance of the model is investigated though, there is a need for a method to select each of the parameters. One approach to adopt the range parameter of the *Markovian* model to the connectivity model of *Henk Wymeersch* is by setting R with a with respect to the proposrtional relationship in beetween the two paramters expressed in (4.50).

We are seeking to investigate whether the two approaches can follow a similar pattern on time with respect to the *Jaccard* index. Thus once the method upon the selection of the range parameter is defined, the next step of our approach is to develop a method for selecting α so that the output of the *Markovian* model meets the connectivity model of *Henk Wymeersch*. To do so we choose to develop a method such that the first component of the *Jaccard* index of *Markovian* model is similar to the corresponding component of the connectivity model of *Henk Wymeersch* denoted as J_{12}. It is expected that the following time instances will follow the same pattern. Derived on this α is estimated as follows:

$$\prod_{i>j}(1-\alpha)^{|\Delta|_J} \cdot \prod_{a_{e_{ij}}(0)=1} exp(-\frac{-|x_2-x_1|_2^2}{2\cdot R^2}) = J_{12}. \tag{6.4}$$

where:

$$\Delta_j = [a_{e_{ij}}(1) = 1 | a_{e_{ij}}(0) = 1]. \tag{6.5}$$

In order to obtain a better estimation upon J_{12} a number of 100 Monte Carlo simulations is conducted and the equation introduced in (6.4) is rewritten as follows:

$$\prod_{k=1}^{100}\prod_{i>j}(1-\alpha)^{|\Delta_j|_k} \cdot \prod_{a_{e_{ij}}(0)=1} exp(-\frac{-|x_i-x_j|_2^2}{2\cdot R^2}) = \prod_{k=1}^{100} J_{12}(k) \tag{6.6}$$

At each iteration an estimate upon the corresponding equality is obtained by solving the following minimization problem :

$$\begin{aligned} \underset{\alpha}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & 0 \le \alpha \le 1. \end{aligned}$$

Where $f(\alpha) = \prod_{i>j}(1-\alpha)^{|\Delta|_J} \cdot \prod_{a_{e_{ij}}(0)=1} exp(-\frac{-|x_2-x_1|_2^2}{2\cdot R^2}) - J_{12}$.Finally the estimate upon α corresponds to estimate per realization that minimizes the equation introduced in 6.6. Based on the pair of the corresponding estimate a number of 100 Monte Carlo realizations is conducted for the two approaches on the same data sets. The obtained results are illustrated in figure 6.6

The obtained results indicate that an offset is initially introduced in between the two approaches. The corresponding behaviour of the *Markovian* model though converges over time to the behaviour of the connectivity model of *Henk Wymeersch*. Derived on this we can claim that the estimator for α yields the desired outcome on

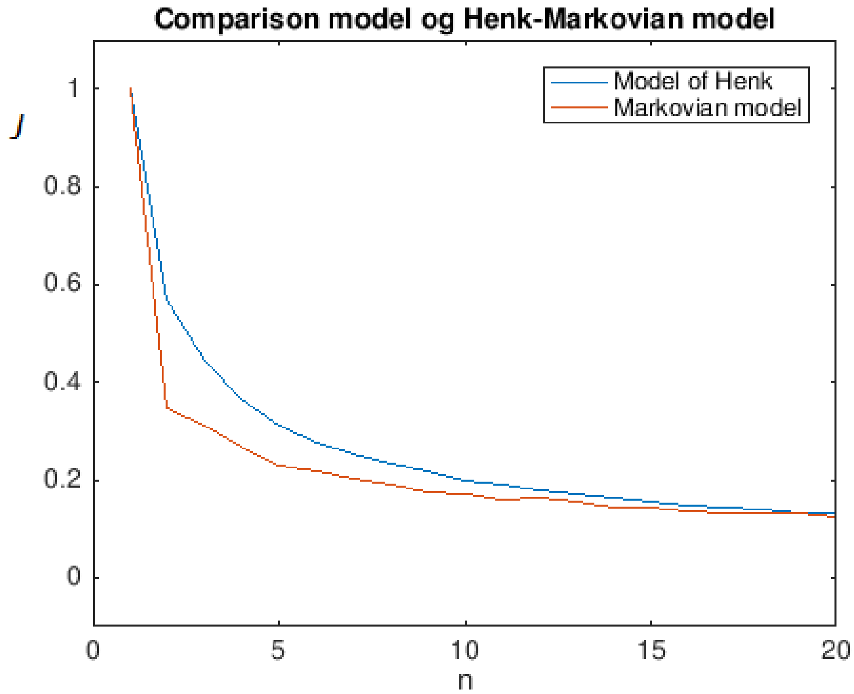


Figure 6.6: A comparison on the performance of the two approaches over time. The x axis represents the time instances whereas the y axis corresponds to the output of the *Jaccard* index.$\epsilon = 0.8$

the *jaccardian* index thus it is a good estimator. We can also argue that the *Markovian* model can adopt the behaviour of the connectivity model of *Henk Wymeersch*. Hence for the cases where the deterministic approach addresses the connectivity issue sufficiently the same argument holds for the *Markovian* model.

6.4 Evaluation of the joint ML estimator

The current work proposed an ML estimator in order to obtain the model parameters. We are considering the GSM technology as a possible implementation for the model therefore we will evaluate the performance of the estimator with respect to the particular criterion. We introduced the results from 100 Monte Carlo simulations in section 5.1.2 the results from ML function for each of the realizations along with the overall evaluation. Note that the simulations took place for 300 different time instances ,due to the heavy computational workload, even though it takes more time for the distribution to become stationary. The corresponding likelihood functions are independent with each other thus they can from the overall likelihood function of the simulations. The joint ML estimator over the 100 Monte Carlo simulations reads:

$$\Lambda(\alpha, R|A) = \prod_{k=1}^{100} \prod_{i>j} \prod_{n=0}^{N} P(a_{e_{ij}}(n) = a_{e_{ij}}^{(d,k)}(n) | (x_i(n), x_j(n))^{(d,k)}, \alpha, R) \tag{6.7}$$

The pair of estimates obtained by the individual likelihoods that maximizes the overall likelihood is the estimate of the overall joint ML estimator. The true pair of values along with the pair of estimates are as follows:

Table 6.2

R_t	$\hat{R}$	α_t	$\hat{\alpha}$
4.57	4.47	$1.55 \cdot 10^{-4}$	$1.98 \cdot 10^{-4}$

In order to evaluate the the impact of the biased estimation over time a number of 100 Monte Carlo simulations. The trivial midrange distance connectivity scenario is considered for the simulations. The obtained results are illustrated in figure 6.7.

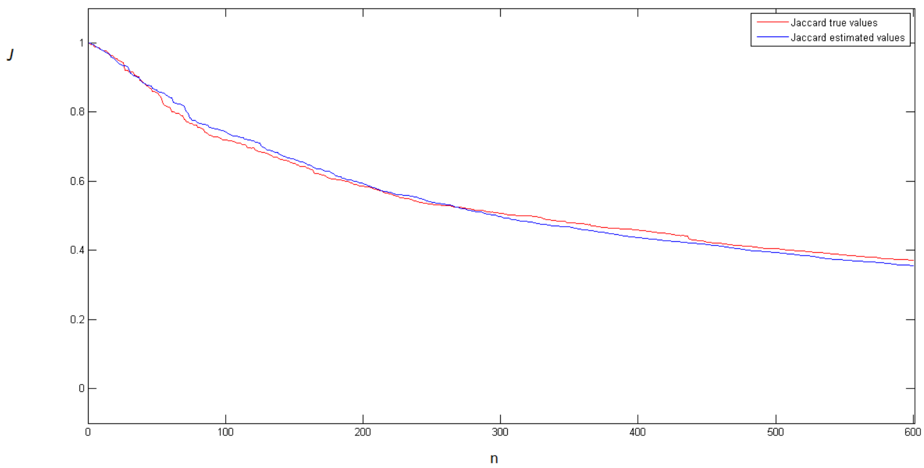


Figure 6.7: A comparison on the performance of the two approaches over time. The x axis represents the time instances whereas the y axis corresponds to the output of the *Jaccard* index.

The obtained results indicate that pair of estimates capture the behaviour of the pair of true values. Thus we can conclude that the joint ML estimator provides a good estimate for the problem at hand.

Chapter 7

Conclusion and Outlook

The primary objective of the thesis has been to propose a connectivity model ,on a dynamic network, for localization purposes. The already existing approaches on the field, a probabilistic model of *Savic* and *Zazo* and the connectivity model of *Henk Wymeersch* of *Henk Wymeersch*, address different aspects of the issue such as positions of the agents or impact of the surroundings. A key aspect that is not taken into account in much detail though has to do with time. Therefore the main focus of the current work is on this aspect and the way that connectivity varies due to this aspect. Hence we introduce a *Markovian* model to address the connectivity issue which is a generalization of the original proposal by *Savic* and *Zazo* in the time domain.

The already existing methods even though they address the same problem they are quite different with each other. Thus a comparison in between them considering a dynamic network is difficult to be done straight forward i.e topologically. The *Markovian* model passes the initial approach to the time domain thus one way to compare the different approaches is in the particular domain. A dynamic connectivity network can be graphically represented by a graph. There is not any method to compare graphs in the time domain. Therefore this thesis proposes the *Jaccard* index in order to investigate how the connectivity of a graph evolves over time. The Jaccard metric is a computationally efficient method upon we can investigate the evolution of a dynamic graph over time. On the other hand it is a hard to derive an analytical expression of the metric so we can have an expectation regarding the outcome and evaluate it.

The model proposed by the current thesis consists of two parameters which makes model more flexible and thus adoptive in different conditions. The comparison in between the *Markovian* model and rule of *Henk Wymeerch* indicated that the proposed method can capture the behaviour of the existing method. Hence we can claim that the proposed approach can cover more aspects of the connectivity problem compared to the existing method.

Several issues can be further considered about the model. One of the main consideration takes place upon the expression of the *alpha* parameter. The dynamic behaviour of the environment alters the signal propagation environment thus by making use of a static parameter to express the probabilist for loosing a connection. A future direction upon this issue could consider a time dynamic variation of the parameter with respect to the interpolated distance in between the agents. Another consideration upon the connectivity model is that uses a particular approach to address each of the other modelling aspects of the Cooperative Localization paradigm regarding the dynamic behaviour of the agents and the noise introduced in the measurements. There exist many different mobile entities involving in a dynamic network, i.e vehicles, bicycles, human beings etc. The corresponding mobility model may not capture the characteristics of the different participants so different mobility models may be required to address the issue per node. The range error model followed by the current work may also not be efficient for the signal propagation environment since many different factors may occur which vary from one area to the other. The current task of this thesis was to introduce a new approach to address the connectivity issue and validate the functionality of the method. Hence we followed a simple approach to address different aspects of the connectivity issue. A possible application of the model when considering a specific region could take these aspects into account. The main scope of the current work is to propose a method upon such an application can be based.

The *Markovian* model is dual parametric. On one hand the flexibility of the model makes it adoptive in different conditions but on the other hand we need a method upon model calibration can be based. Therefore the current work proposes an ML estimator in order to obtain the model parameters. The obtained results indicated that the ML estimator is a good method for the calibration purposes when considering the GSM technology as a possible application of the model. The obtained results indicated that there is a dense presence of outliers on the estimates. This means that the required data to address the necessities of the estimation task for the model. There exist tools, such as the Cramer–Rao bound that can be used to exploit information upon this matter.

One of the aspects that need to be addressed in the future is the behaviour of the *Markovian* model when it is adopted by a possible application such as a Cooperative Localization algorithm i.e the Variational Message Passing Algorithm [10]. A study upon the behaviour of the model with respect to the corresponding application could lead to a development of an application which is based upon the structure of the *Markovian* model.

Bibliography

[1] B. Çakmak et al. "Cooperative Localization for Mobile Networks: A Distributed Belief Propagation - Mean Field Message Passing Algorithm". In: *IEEE Signal Processing Letters* 23.6 (2016), pp. 828–832. ISSN: 1070-9908. DOI: 10.1109/LSP.2016.2550534.

[2] Emre Atsan and Öznur Özkasap. "A classification and performance comparison of mobility models for ad hoc networks". In: *Ad-Hoc, Mobile, and Wireless Networks*. Springer, 2006, pp. 444–457.

[3] Peter Bull et al. "A multi-class mobility model for dynamic and dependable systems". In: *Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on*. IEEE. 2013, pp. 1010–1015.

[4] Charles S Carrano, Keith M Groves, and James M Griffin. "Empirical characterization and modeling of GPS positioning errors due to ionospheric scintillation". In: *Proceedings of the Ionospheric Effects Symposium, Alexandria, VA*. 2005.

[5] Ann Cavoukian and Kim Cameron. *Wi-Fi Positioning Systems: Beware of Unintended Consequences Issues Involving the Unforeseen Uses of Pre-existing Architecture*. Information and Privacy Commissioner of Ontario, Canada, 2011.

[6] Ahmed Helmy FanBai. "A survey of mobility models in wireless Ad Hoc Networks". In: *University of Southern California, USA* ().

[7] Geoffrey Grimmett and David Stirzaker. *Probability and random processes*. Oxford university press, 2001.

[8] X Rong Li and Vesselin P Jilkov. "Survey of maneuvering target tracking. Part I. Dynamic models". In: *Aerospace and Electronic Systems, IEEE Transactions on* 39.4 (2003), pp. 1333–1364.

[9] Singthongchai Niwattanakul et al. "Using of Jaccard coefficient for keywords similarity". In: *Proceedings of the International MultiConference of Engineers and Computer Scientists*. Vol. 1. 2013, p. 6.

[10] C. Pedersen, T. Pedersen, and B. H. Fleury. "A variational message passing algorithm for sensor self-localization in wireless networks". In: (2011), pp. 2158–2162. ISSN: 2157-8095. DOI: 10.1109/ISIT.2011.6033940.

[11] Troels Pedersen and Carles Navarro Manchon. "Lecture Notes in Stochastic Processes".

[12] Vladimir Savic and Santiago Zazo. "Sensor localization using nonparametric generalized belief propagation in network with loops". In: *Information Fusion, 2009. FUSION'09. 12th International Conference on*. IEEE. 2009, pp. 1966–1973.

[13] Richard Weber. *Lecture Notes On Markov Chains*. http://www.statslab.cam.ac.uk/~rrw1/markov/M.pdf.

[14] H. Wymeersch, J. Lien, and M. Z. Win. "Cooperative Localization in Wireless Networks". In: *Proceedings of the IEEE* 97.2 (2009), pp. 427–450. ISSN: 0018-9219. DOI: 10.1109/JPROC.2008.2008853.

[15] Reza Zekavat and R Michael Buehrer. *Handbook of position location: Theory, practice and advances*. Vol. 27. John Wiley & Sons, 2011.