OULUN YLIOPISTO
UNIVERSITY of OULU
DEGREE PROGRAMME IN WIRELESS COMMUNICATIONS ENGINEERING

MASTER'S THESIS

DECISION ERROR PROBABILITY IN A TWO-STAGE COMMUNICATION NETWORK FOR SMART GRIDS WITH IMPERFECT SENSING AND DATA LINKS

Author	Iran Ramezanipour
Supervisor	Doctor. Pedro Nardelli
Second Supervisor	Docent. Premanandana Rajatheva

December, 2015

Ramezanipour A. (2015) DECISION ERROR PROBABILITY IN A TWO-STAGE COMMUNICATION NETWORK FOR SMART GRIDS WITH IMPERFECT SENSING AND DATA LINKS. University of Oulu, Department of Communications Engineering, Master's Degree Program in Wireless Communications Engineering. Master's thesis, 62 p.

ABSTRACT

This thesis analyzes a scenario where the distribution system operator needs to estimate whether the average power demand in a given period is above a predetermined threshold using a 1-bit memoryless scheme. Specifically, individual smart-meters periodically monitor the average power demand of their respective households to inform the system operator if it is above a predetermined level using only a 1-bit signal. The communication link between the meters and the operator occurs in two hops and is modeled as binary symmetric channels. The first hop connects individual smart meters to their corresponding aggregator, while the second connects different aggregators to the system operator. In the first set of analysis, the decision making only happens by the network operator in the second hop and aggregators in the first hop only work as relay nodes which only forward the information it has received from the smart meters. AND and OR decision rules are studied in this scenario. Moreover, in the second set of analysis, the decision about the power demand happens in two stages based on the received information bit. Meaning that the decision making happens both by the aggregators in the first hop and network operator in the second hop. We consider here three decision rules in the second scenario: AND, OR and MAJORITY. Our analytical results indicate the circumstances (i.e. how frequent the meters experience the consumption above the defined threshold) and the design setting (i.e. decision rules) that a low error probability can be attained. We illustrate our approach with both theoretical and numerical results from actual daily consumptions from 12 households and 3 aggregators. Also, we derive closed-form equations for the average decision error probability as a function of the system parameters (e.g. number of sensors, communication error, sensing error) and the input signal characterization. The first set of simulations are done in Matlab. Since the second set of data are provided in Excel; thus, the simulations are done using Visual Basic.

Keywords: Decision theory, communication networks, error probability, smart grids, wireless sensor network

TABLE OF CONTENTS

FOREWORD

This thesis work was carried out at the center for Wireless communications of university of Oulu, Finland. The aim of this thesis work is to design a communication system for smart grids using a wireless sensor network with low cost which is also easy to implement and has a low average error probability. The proposed model uses different logical hard decision rules in different stages of the network to decide on the state of the system. The target is to see how practical is this model to be implemented as a way of communication in smart grid networks.

Here, I would like to gratefully acknowledge my supervisor Doctor Pedro Nardelli for the guidance and help he provided during this thesis work. I would also like to thank my second supervisor Docent Premanandana Rajatheva. Finally, I want to thank my family for always believing in me and supporting me throughout the years.

LIST OF ABBREVIATIONS AND SYMBOLS

$\mathcal{N}_i$	Set of Smart meters
i	Consumers (Prosumers)
N_i	Elements of i
N	Number of sensor (users or smart meters)
$\mathcal{N}$	Set of Aggregators
j	Number of hops
t_n	Time instants
γ	Predefined threshold of the Network
$\theta_{i,j}[n]$	Binary function which indicates the meter state
$P_{i,j}(t_n)$	Average power demand at time instant t_n
q	Sensing error
p_1	NAN communication error
p_2	WAN communication error
$\theta_i[n]$	Binary state of the aggregators
$\theta[n]$	Global binary state of the system
$x(t)$	Input signal
$y[n]$	Aggregators Input (sensors output)
$\hat{\theta}[n]$	Global estimated binary state of the system
$s_i[n]$	First hop output
$\mathcal{E}$	The event under observation
$S_{i,0}[n]$	Signal sent by sensor i
$S_{i,j}[n]$	State of the system at jth level
j	Number of hops
$\theta_{i,0}[n]$	Initial binary state of the network
$g(y_N[n])$	Bolean function that estimates $\hat{\theta}[n]$
$\bar{P}_{i,j}(t_n)$	Average power demand of the whole network
$n_{\max}$	Number of measurements considered
P_e	Average error probability
$P1-0$	Probability of receiving zero when the input is one
$P0-1$	Probability of receiving one when the input is zero
$\hat{s}$	received signal
Pr	Probability
H	House
K	Number of aggregators

AES	Advanced Encryption Standard
AMI	Automatic metering infrastructure
AMR	Automatic meter reading
BAS	Building automation system
BSC	Binary symmetric channel
CDMA	Code division multiple access
CWC	Center for wireless communications
DES	Data encryption standard

DLC	Direct load control
DMS	Distribution management system
DSSS	Direct sequence spread spectrum
EMS	Energy management system
ESI	Energy services interface
FFD	Full function device
GMSK	Gaussian minimum shift keying
GPS	Global positioning system
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HAN	Home Area Netwrok
IED	Intelligent electronic device
ISO	Independent system operator
IEEE	Institute of electrical and electronics engineers
LAN	Local Area network
LMS	Load management system
LOS	Line of sight
LTE	Long term evolution
MIMO	Multiple-Input and Multiple-Output
MAC	Media Access Control
MV	Medium voltage
NAN	Neighborhood Area Network
NLOS	Non line of sight
NIST	National Institute for Standards and Technology
OFDMA	Orthogonal Frequency-Division Multiple Access
OQPSK	Offset quadrature phase-shift keying
PLC	Power Line Communication
REDD	Reference Energy Disaggregation Data Set
RFD	Reduced function device
SEP	Smart Energy Profile
WAN	Wide area network
UMTS	Universal mobile telecommunications system
WCDMA	Wideband code division multiple access
WiMAX	Worldwide Interoperability for Microwave Access
WSN	Wireless sesnsor networks
2G	2nd generation
3G	3rd generation

1. INTRODUCTION

The basic structure of the electrical power grids have been the same for the past century and even though they have been contributing greatly to our daily and industrial needs, they are not suitable for the needs of the modern days any more. Several problems such as more voltage sags, blackouts, and overloads have been raised from using the traditional power grid systems, especially in the past decade as a result of slow response time of devices over the grid [1]. In addition to that, with the growing size of the population, the demand for electricity and consumption is also increasing [2]. This means more appliances and consumers are joining the current power grids which are not designed for handling these large amounts of users. Also, the current power grids contribute greatly to the carbon emission [3]. With both economic and environmental aspects in mind, changing such an unstable and inefficient system into a more interoperable, secure, and cost effective systems seems inevitable. [1].

The concept of smart grids has been introduced to characterize the modernization of the traditional electrical power grid, empowered by the advances of information and communciation technologie [4]. Smart grid is a modern power grid system with improved efficiency, reliability and safety [2]. This new power grid is called Smart grid which is likely to be a very good solution to the current energy crisis [5] and this change will definitely affect utilities, regulation entities, service providers, technology suppliers, and electricity consumers too [6]. Since the smart grid is based on a two way, almost real time communication between different elements of the network [7], communication technologies is one of the most important part of a smart grid [8, 9]. Communication technologies have helped smart grids become more achievable [10].

In this thesis, we are proposing a new structure for the communication systems in a smart grid network using a wireless sensor network (WSN) [11], since the wireless sensor networks are becoming more popular each day to be used in any possible application area [12]. Our study was done on general applications with loose reliability requirements. There have previously been some researches done in implementing WSNs in smart grids networks (e.g. [13, 14]). In our research, we are trying to build an efficient communication system where its low cost and simplicity are its main characteristics. With the help of different decision rules and logical gates, we created a communication system based on WSNs which can detect a given event based on an input signal, keeping its decision as simple as possible. What makes the concept studied in this thesis different form previous works is that we have studied the efficiency of using the one bit signaling method in detecting an input signal based on the decisions that have been made through out the network by the aggregators and network operators.

The design and implementation of communication networks in smart grids face several challenges. The main problem is that there are no specific communication infrastructures that have been widely accepted in order to transform the current power grids into smart grids [1, 15, 16] and so several wired and wireless communication technologies such as Power line communications [17], optical fibers [18], IEEE 802.11 based wireless LAN, IEEE 802.16 based WiMAX,3G/4G cellular, ZigBee based on IEEE 802.15, are currently being used in Smart grids communication technologies [19].

Due to the different existing applications in such large infrastructure (smart grids), the requirements for the communication systems are diverse and context-dependent. Some control applications should be almost real-time and highly robust while household billing has looser needs [20].

Moreover, in the first part of this research, we will see the average error probability of the system in a two hop system model where the decision making is down at the end point of the second hop. Smart meters which are operating with sensing error send their data to aggregataros through a communication channel where the data faces the first communication error, then this data is sent by the aggregators to the network operator through a new communication channel where again the data meets another communication error, finally, the network operator decides on the overall state of the system based on the information it has received. Different error probability equations are derived for different decision rules and different situations are simulated in order to study the system behaviour.

In chapter 4, we consider the case where the smart meters need to inform the aggregator whether their energy consumption (or generation, or the balance between them) in predetermined time periods is above or below a given threshold. Aggregators proceed similarly with the system operator based on individual meter information. The system operator then needs to decide about the general state of the whole network. Our proposed model consists of two layers, first layer is the connection between the houses and network aggregators. There are a total of 12 houses studied in this chapter. We have grouped them in 3 different networks consisting of 4 houses each. Each of these networks sends its data to its corresponding aggregator, the aggregator then decides about the state of the first layer and send it to the second layer which is defined as the connection between the aggregators and the network operator. Such information obtained from the proposed network can then be used in, for example, cloud computing applications [21] such as forecasting, demand-side management, peer-to-peer energy trading.

2. COMMUNICATION SYSTEMS FOR SMART GRIDS

The traditional power grids that are being used currently to transmit power from the utilities to the consumers have been in use for more than a hundred years. Technology is advancing with a fast speed, not to mention the fact that the population of the world is also growing which means that more users are being added to the power grids around the world. In addition the mentioned facts, the current power grids main resource are fossil fuels, this would mean that our current power system is responsible for carbon emissions and pollution. Considering all the mentioned facts, it is easy to understand why the old power grids are not suitable for todays needs anymore and why the are facing several problems such as more blackouts and voltage sags.

This means that the current power grids need to be replaced by an alternative that is both reliable and efficient. This new structure should be able to address the problems that were mentioned earlier in addition to being scalable, manageable secure and cost effective. This new structure with all these characteristics is called Smart grids [1].

Smart grids can be defined as the integration of the modern communication and information technologies with the traditional power grids which will result in an improved efficiency,reliability and safety. Also, smart grid do not depend on fossil fuels and can use renewable energy sources; thus, decreasing the carbon emission. Using smart grids will result in an enhanced system throughput and also will reduce the consumption on the grid [2].

A power grid consists of many different parts such as utilities, transmission lines and consumers. Smart grids too consist of different elements such as intelligent electric devices (IED), sensors whether wired or wireless, smart meters, distributed generators and dispersed loads. These different elements of the grids need to be able to cooperate with each other in order for the system to work as expected. This cooperation is done using different types of communication and information technology. Therefore, it can be said that the communication system is one of the most important parts in a smart grid network [6].

As it was mentioned earlier, what changes the power grid into a smart grid is adding a communication network to the electrical grid with the main responsibility of collecting and analyzing different aspects important to the grid operator. The functionality of a smart grids highly depends on this real time information that it receives from different parts of the network and it is based on this real time data that the smart grid can act on managing power. This definition alone can show how important is the communication network to the smart grid; thus, the design and implementation of the communication network has a big impact on the grid as it is the tool that makes the communication between different parts of the grid possible [1].

Fig1 shows the basic concept of a smart grids which may help for better understanding the reason why the current power grids are being replaced by the smart grids.

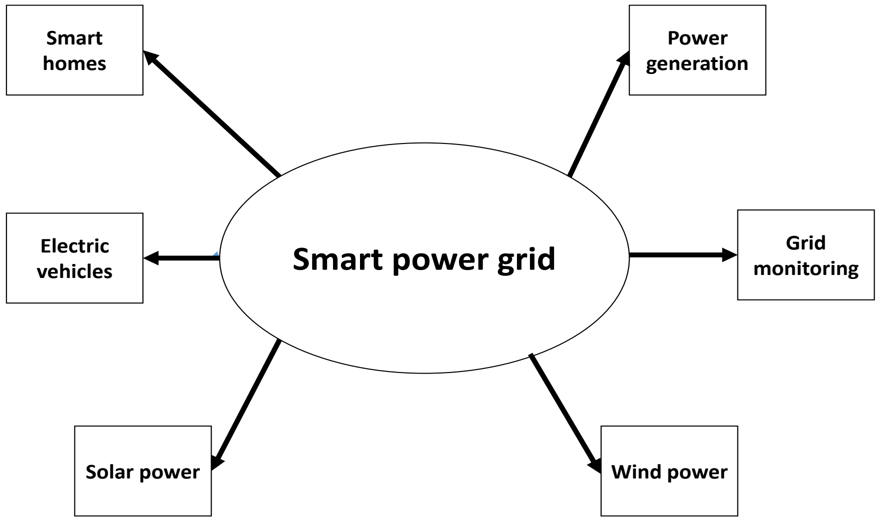

Figure 1: Basic illustrarion of a smart grid .

Although communication network is very important to a smart grid, it also faces several challenges. The most important problem can probably be the fact that the requirements and specifications of a communication system in a smart grid is not very straightforward and therefore there is no specific architecture that would be accepted by the majority of companies. That is why different companies usually design there own communication structure for being implemented in a smart grid [1].

Before we start looking at some of the most popular communication technologies that are currently being used in smart grids networks, let us introduce some of the main characteristics of a communication system which is being used in a smart grid in addition to reliability, extensibility, scalability, manageability and low latency [8].

- high availability,
- automatic management of redundancies,
- appropriate communication delay and system responsiveness,
- high security,
- ease of deployment and maintenance.

Moreover, there are two types of subsystems that are used in the communication systems in smart grids. First one is a communication infrastructure which should be able to carry out the tasks that is expected from the communication system of a smart grid and most importantly keep the connection between different elements of the grid alive at all times. This infrastructure is a combination of communication technologies, networks and protocols. The second subsystem is a middleware platform. This platform is where the software layer which is the communication tool between the applications and the first subsystem exists [18].

Further in this thesis we will see different types of smart grids communication networks that are used in out analysis, for better understanding each of them, here are a short explanation of each [1]:

- Home Area Networks (HAN) which as can be understood from its name is a communication network that connects different house appliances. These appliances all have smart meters which is responsible for monitoring various grid elements such as gas, electricity or heat.
- Neighborhood Area Networks (NAN) which can be explained as a network which consists of several HANs connected together
- Wide Area Networks (WAN) which is the larges network in an smart grids and includes both HAN and NAN. WAN is what makes the two way communication with the network operator possible.

As it was mentioned earlier, there is no specific communication system for being used in smart grids and there are several different technologies that are being implemented currently in smart grids. Table 1 shows some of these most popular communication technologies and their properties [22].

Table 1: Smart grid communication technologies

Technology	Spectrum	Data rate	Coverage range	Application	Limitation
GSM	900-1800 MHz	Up to 14.4 Kpbs	1-10 Km	AMI, Demand response, HAN	Low data rates
ZigBee	2.4 GHz-868-915 MHz	250 Kbps	30-50 m	AMI-HAN	Low data rate,short range
PLC	1-30 MHz	2-3 Mbps	1-3 Km	AMI, Fraud Detection	Harch, Noisy channel environment
3G	1.92-1.98 GHz/2.11-2.17 GHz	384 Kbps- 2 Mbps	1-10 Km	AMI, Demand response, HAN	Costly expectrum fees
WiMAX	2.5 GHz, 3.5 GHz, 5.8 GHz	Up to 75 Mbps	10-50 Km (LOS)/1-5 Km (NLOS)	1AMI, Demand response	Not widespread
GPRS	900-1800 MHz	Up to 170 Kpbs	1-10 Km	AMI, Demand response, HAN	Low data rates

2.1. ZigBee

ZigBee is considered to be the best communication technology for being used in home area networks by the U.S National Institute for Standards and Technology (NIST) since it has many advantages such as having low cost of implementation, being efficient in term of power usage and data rate. The mentioned properties makes ZigBee technology a suitable option for smart grids applications within HANs such as smart lightening, energy monitoring and energy management. That is why a lot of companies dealing with AMI prefers smart meters that are compatible with ZigBee Smart Energy Profile (SEP) [10, 22].

ZigBee was first introduced to the technology by ZigBee Alliances. It is based on Physical layer and Media access control (MAC) layer of the IEEE 802.15.4 standards.ZigBee uses direct sequence spread spectrum (DSSS) and its operating frequency band is ISM bands of 868 MHz, 915 MHz and 2.4 GHz as it was mention in Table.1. Because of its low data rate, ZigBee is a suitable option for home automation applications [10, 22].

There are two different kinds of devices that exists in the networks that use ZigBee technology. One is called a full function device (FFD). Formation, management and data routing are the responsibility areas of the FFD. The other kind of device used in ZigBee enabled networks is called Reduced function device (RFD) which act as the supporter of FFD [10, 22].

One of the main factors to be considered in a smart grid network in the security of the network. ZigBee benefits from a strong authentication process between the devices that want tp communicate with each other in the network. This authentication process is based on 128 bit AES encryption [10].

ZigBee is a suitable option which makes the wireless networking between the different elements of the grid possible. It has 16 channels in the 2.4 GHz band and each of these channels has a bandwidth of 5 MHz and uses Offset quadrature phase-shift keying (OQPSK) modulation method. On the down side, ZigBee has some disadvantages such as low processing capabilities, small size of the memory and small delay requirements. It also suffers from interference between the devices.

2.2. WiMAX (IEEE-802.16)

Another communication technology that is currently being used in the smart grids systems is the Worldwide Interoperability for Microwave Access (WiMAX) under the IEEE-802.16 standards for Wireless broadband. In this explanations, IEEE-802.16 indicates the physical and MAC layer of the WiMAX technology. This physical layer uses OFDMA modulation scheme in addition to other capabilities such an antenna system which is designed using the MIMO technique. Using this method enables the WiMAX technology to use LOS on its $11 - 66$ GHz frequency band and NLOS on its $2 - 11$ GHz frequency band.

In terms of security, WiMAX uses data encryption standard (DES) and also AES encryption techniques like Zigbee and therefore has a secure communication. The fact the main purpose of the WiMAX design is to be used in point to multi point communication in addition to other capabilities such as long communication range ,low cost of implementation and high data rates, makes the WiMAX technology a good backbone for a smart grid network [23].

2.3. Cellular network communications (GSM, GPRS, 3G)

One of the suitable options for communication systems to be used in smart grids is our current cellular system such as 2G, 3G, WiMAX and LTE. Since they are already designed, implemented and are in use, using the as a communication tool between the smart meters and the utility and other network elements would be very cost effective for companies since they would not have to spend extra time and money on designing and implementing a new communication infrastructure which would be suitable for their specific application.

Assuming that smart meters in a network send their data to the corresponding utility every 15 minutes, this indicates that every 15 minutes there a great amount of data that has been generated and needs to be send through the communication channel. Handling a big amount of data requires a high data rate connection, that is why many telecommunication companies such as T-Mobile, Vodafone and Telenor have agreed to make it possible for the smart grid networks to use their GSM network for their data flow from the smart meters to the utility. In addition to this, some of the smart meters in the industry are compatible with the GPRS technology. Other wireless communications technologies that are currently being used in some smart meters project are code division multiple access (CDMA), Wideband code devision multiple access (WCDMA) and Universal mobile telecommunications system (UMTS) [22].

GSM is the most used cellular network communication technology in smart grids. Its operating frequency bands are 900 MHz and 1800 MHz. Gaussian minimum shift keying (GMSK) is the modulation method that is used in GSM and. GSM has a high level of security and is most popular with application in HAN, home monitoring and load control [10].

Our current cellular network have a lot of advantages that makes them the best option for being used as the communication technologies in smart grids. Some of these advantages are low cost of implementation and maintenance, secure data transmission and fast installation. However, as we all have experienced situations that there is no cellular coverage. The fact that cellular communications might not always be available since a lot of users are always connected to it which will result is congestion or decrease in network performance, will result in companies that are dealing with critical application not to chose the current cellular communication technologies and build their own communication infrastructure which is suitable for their needs [22].

2.4. Power Line Communication (PLC)

As its name suggests, smart grids networks that use PLC as their communication system use the existing power lines in order to transmit data between the devices in the network with a relatively fast speed (2-3 Mb/s). Since this type of communication is a wired communication technology and not wireless, it is the most popular choice of communication when it comes to the electricity metering since the wired connection is directly connected to the meter. PLC networks are usually implemented in a way that the power lines are the mean to connect the smart meters to the data collector and then information is sent to the data center using the cellular communication technologies mentioned earlier. In other words, PLC is the communication mean between the smart meters and data concentrator and GPRS is tool of communication between the data concentrator and the utility. One of the pints that make PLC a popular choice of communication is that any electrical devices which are equipt with smart meters can be connected to the power line. That is why in some countries like France, companies are upgrading the traditional meters with PLC compatible meters.

As we know, smart grid is about the two way communication between the elements of the network and the utility but the current traditional power lines are incapable of this two way communication. That is why they have to be equipt with a modulated carrier signal over them that makes this two way communication possible.

PlC has two different categories. One of them is the Narrowband PLC which operates on the 3-500 Hz frequency range and is usually used in sensing and communication applications in smart grid networks. The other category is the Broadband PLC. Broadband PLC operates on the 2-250 MHz frequency range and in addition to its smart grids applications, it is also used in internet applications and end user entertainment. PLC is currently being used in various applications of smart grids and therefore, it is considered as one of the most popular communication systems in smart grids. PLC based smart meters, automation of the Medium voltage (MV) grids and advanced metering infrastructure (AMI)are some of the main applications of the PLC systems.

Although Power Line communications have become very popular for being used in smart grids, it also suffers from some disadvantages too. For example, PLC is yet not suitable for HANs. What makes PLC unsuccessful when it comes to HAN environments is the lack of PLC standards which can be adaptable with other communication technologies. Another problem with PLC technology is that it suffers from additive non-Gaussian noise [1, 2, 8].

3. SYSTEM MODEL AND PROBLEM FORMULATION

3.1. Problem Formulation

Since the traditional power grids are not suitable for the needs of the 21st century thus they are being replaced by smart grids networks in recent years . Intelligent electronic devices (IEDs), wired and wireless sensors, smart meters, distributed generators and dispersed loads are what form a smart grid networks. The cooperation and coordination of these components rely on the communication network of the smart grids; hence, information and communication systems are considered to be the most important part of a smart grid. In other words, the difference between the traditional power grids and smart grids is the integration of the communication networks with the electrical grids and capturing and analyzing the real time data it receives from the power grid.

It has been previously mentioned that there is no specific communication network structure for the current smart grids and so companies usually propose their own communication network structures in order to be used for different applications. In this thesis we are proposing a new way of decision making in a smart grid network in an attempt to minimize the error in the final data which is received by the network operator.

In our general proposed model, we assume a network of houses which are connected together in a star network topology. The system model is shown in Fig.2.

This network consists of different layers. In the first layer we have the devices in the houses. each house includes several different devices, each of them having a sensor (smart meter). For instance, in case of house number one, we assumed three devices with three sensors, S11,S12 and S13. Each of these Sensors measure and store the data based on their application, whether it is to measure and store the temperature, voltage usage, power usage, etc. As can be seen in Fig.2, all the smart meters in one house are connected to a network aggregtor (e.g.smart meter 1 in case of house one).

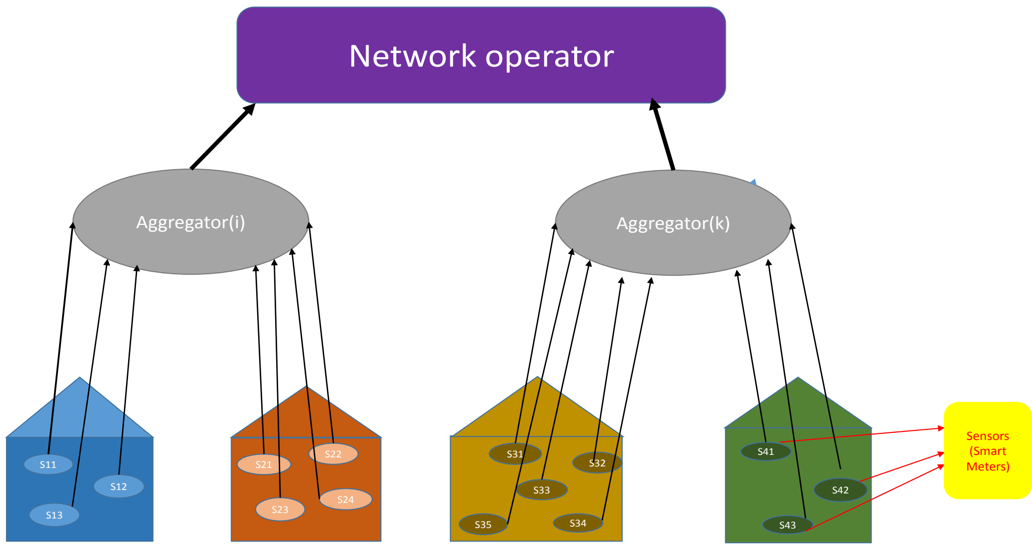

Figure 2: System Model

In this part, we consider a network where smart meters in the houses are programmed to monitor if a given event has happened in the network in order to be able to inform the network operator whether this event happened or not so that the network operator can take the suitable reaction. We study the probability of a decision error happening in the network.

Sensors, smart meters and the connection links between the smart meters and the network operator are prone to errors in this scenario, whether it is sensing error or communication error.

The problem studied in this thesis consists of two parts. The first part is the theoretical evaluation of the proposed model which is presented in this chapter. The next chapter will be about the practical implementation of the model studied in this chapter.

The general system model in this thesis is consist of two different layers. The first layer consist of the smart meters in the houses which are connected to the aggregators. These smart meters send their measured data to the aggregators. These aggregators are the end point of the first layer of this network. In this chapter, the first fist hop of the network (aggregators) works as a relay network which means that no decision making happens in the first layer. So when the data sent by the smart meters reaches the network aggregators, it is sent to the next hop which is the network operator using an error prone communication channel without any implementation of the decision rules. In this chapter, the only place that the decision making happens is the last hop by the network operator. In the theoretical results presented in the next chapter, network aggregators have to make a decision about the state of the network based on the individual meter information. After that, the aggregators send their decision (state of the first layer) to the second layer of the network which would be the system operator. The system operator then makes a decision based on the data it has received from the aggregators on about the state of the whole network. This would be explained in details in he next chapter.

The above mentioned decisions are based on hard decisions made by different logical gates. In this thesis, the two most popular logical decision rule which are AND and OR rules that are studied in the theoretical part.

In this chapter, we derive the error probability of each layer of the network and based on that we define the best strategy to be used in different situations. The effect of different network parameters are also studied. For instance, the effect of increasing the number of meters (users) in the network, or how having a reliable and non reliable network can affect the error probability and reception of the correct data. In our proposed model, two kinds of errors are considered, the sensing error which is caused by the sensors and the communication error which exists in the communication channel of the network.

What we are trying to do in this thesis is trying to understand how a wireless sensor network can be constructed in order to efficiently detect a given event based on the input signal. The simplicity of this sensor network is also very important is this study. An efficient implementation of such networks in smart grids can have a big impact on the current smart grid communication systems in case of efficiency, accuracy and robustness. As we know, some control applications need to be almost real-time and highly robust while some other applications such as household billing has looser needs.

The results of this thesis can be used to define the suitable structure for each of these applications.

3.2. System model

In this part, we present a framework to evaluate the probability that a decision error event occurs in wireless sensor networks. This analysis includes both sensing and communication errors. As it was previously mentioned, sensors in the first hop need to identify whether a given event has occurred based on its periodic, noisy observation of the signal. This information then needs to be sent to the system operator where it decides about the general state of the network in a given time period. The mentioned information are transmitted from the sensors to the network operator through a binary symmetric channel which have different error probabilities. The decision that is made by the network operator is based on different logical operations such as AND and OR which are studied in this chapter in addition to MAJORITY which is studied in the next chapter. It is shown that the average error probability equations derived in this thesis are a function of different network parameters such as number of sensors, number of hops and sensing and communication error.

In this research we are following a different approach compared to what have been done before (e.g. [24–32]). What makes this research different is that here the quantization and decision rules are given in the scenario under analysis whereas in the other mentioned works, the focus is on finding the optimal detection/estimation scheme. Our approach does not focus on high reliability or low latency, but rather on a cheap way to estimate the average power demand without harming the communication network with huge amounts of data (e.g. [33]).

As it was previously mentioned, our aim in this thesis is to evaluate the average error probability of our proposed WSN with little computational capabilities. Specifically in this chapter, out goal is to evaluate a scenario where a set of sensors periodically measure a given signal to detect whether or not a given event has happened (e.g. if the level of the signal is above a given predefined threshold or not or the temperature of their respective site is above or below a given temperature). Based on their noisy measurements, the occurrence of such event is then converted into the binary format where “1" implies that the event has happened and “0" implies that the event has not happened. These binary values also define the sensors state. The sensors then need to send their state to the network operator using a wireless channel which can have one or multiple hops. The sensors first send their data to aggregators, these aggregators work as relay nodes in this chapter, meaning that no decision making is done by them and they only forward the data they have received to the network operator. The communication channels are assumed to be binary symmetric channels and their associated error can be different at each hop. Although these communication error probabilities can be different considering different networks such as NAN and WAN [20], the communication error probability is the same within each of these networks. Network operator which is the last point of the second hop in this model makes a decision about the general state of the network based on the binary states that it has received related to the state of each of the sensors. The decision rules employed by the network operator are the Boolean functions, AND and OR which are studied in this chapter and MAJORITY which is studied in the next chapter in addition to the other two functions.

Let us assume a network composed by a set $\mathcal{N}_i = \{1, ..., N_i\}$ of smart meters [34] of a given group of consumers (prosumers) i composed by N_i elements, which are associated with aggregator $i \in \mathcal{N}$ where $\mathcal{N} = \{1, ..., K\}$ is the set of aggregators. Each meter $j \in \mathcal{N}_i$ needs to inform aggregator i in predetermined times $t_n = t + n\tau$ if its individual observation of the system and if the event is above or below a given threshold γ.

Let $\theta_{i,j}[n]$ denote the binary function that indicates the meter state based on the input signal level at time t_n. We consider $\theta_{i,j}[n] = 1$ with probability $1 - q$ and $\theta_{i,j}[n] = 0$ with probability q. Where q is the metering error of the smart meters.

We assume smart-meter $j \in \mathcal{N}_i$ sends its state $\theta_{i,j}[n]$ to aggregator i through a binary symmetric channel [35, Ch.7] with error probability p_1 (where the subscript "1" refers to the first stage of the communication). Based on the meter information, aggregator $i \in \mathcal{N}$ send its state $\theta_i[n]$ to the network operator based on the data it has received from the smart meters. Note that since the communication links connecting the smart meters in different houses to their corresponding network aggregator are defined as a neighborhood area network(NAN) [20] and the links connecting the network aggregtors to the system operators are defined as wide area networks(WAN) [20], in our analysis, we have assumed that the errors associated with each of these networks are almost the same, so the error probability in the communication links connecting the houses to the aggregators are considered almost the same in all the three NAN networks (p_1). Also, since the links connecting the aggregators to the network operator are all part of the same WAN networks, the links error are considered the same (p_2).

As mentioned previously, aggregators $i \in \mathcal{N}$ then needs to send its state $\theta_i[n]$ to the system operator in a binary symmetric channel with error probability p_2. With the information from all aggregators in hand, the operator similarly proceeds to decide the global state $\theta[n]$ based on AND and OR logic operations. In this model, we assumed a network of houses which are connected together in a star like network topology.

An illustrative figure of the scenario under analysis is shown in Fig.3. It is shown that Sensors monitor a given signal $x(t)$ in order to determine the binary state $\theta[n]$ at time t_n. Each one of the N sensors in the network needs to send its state to the network operator that remotely decides the state $\hat{\theta}[n]$. On its way to the network operator, errors may happen either in sensing ($y_i[n] \neq \theta[n]$ with $i = 1, .., N$) or in communicating ($s_i[n] \neq y_i[n]$). The dashed rectangle identifies where the error events may happen.

In this chapter, In addition to network operator decision function, we are going to analyze different design structures that can be implemented using WSN in order to improve the average error probability; hence, having a better estimation of the input signal at the output of the network operator.

Here is an example for better understanding of Fig.3. let us assume that $x(t)$ represents the temperature of a factory. x_{th} is the temperature threshold which if higher than a given value, indicates that there is a fire happening in the factory and the network operator has to be informed in order to take actions. The event $\mathcal{E}$ can be then associated with an emergency where $x(t_n) > x_{\text{th}}$.

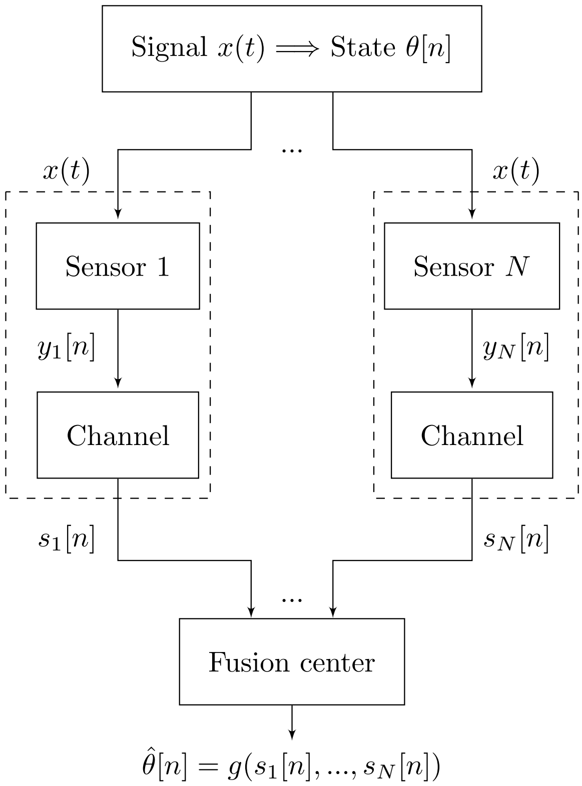

Figure 3: System model studied in this chapter

For better understanding of the communication error, let us consider $S_{i,0}[n]$ to be the signal sent by sensor i and $S_{i,j}[n]$ be the state of the system at jth level. j is defined as the number of hops in the network and can be $j = 1, ...M$, the studied model in this thesis consists of two hops, one from the networks of sensors to the network of aggregators and the other one from the networks aggregators to the network operator. At every hop, the binary state of the network ($\theta_{i,j}[n]$) is forwarded to the next hop using the communication channels which as have been mentioned before, are subjected to error. The probability tree of the state of the sensors $i = 1, ...N$ is presented in Fig.4 (sensing error is not shown in this figure).

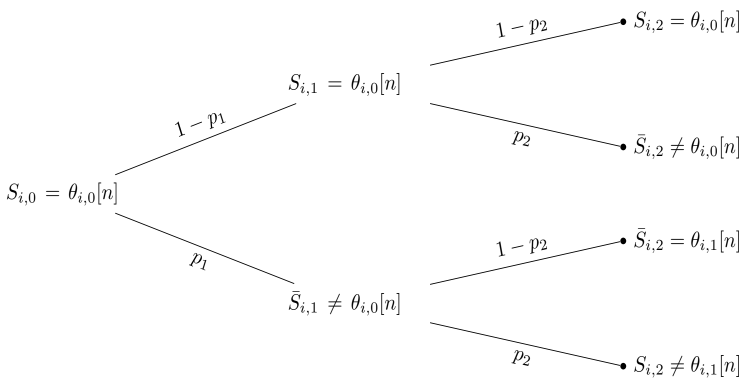

Figure 4: Schematic of the decision making tree in a two layer model

Probability tree of the state of the sensors $i = 1, ...N$ considering the communication error p_j where j indicates the number of hops between the smart meters and the network operator is shown in Fig.4. The initial binary state of the network is $\theta_{i,0}[n]$. Error probability p_j is what causes the output of this channel to be different from the input. The assumption is tht the errors happening in different layers of this model at different time-steps t_n are independent from each other; Hence, the state of the system will be transmitted to the next layer with error probability p_j where j defines the number of the layer (hop) of the system.

This means that the state $S_{i,j} = S_{i,j-1}$ with error probability $1 - p_j$ and $S_{i,j} \neq S_{i,j-1}$ with error probability p_j.

As can be seen, this decision making tree only illustrates the communication error in a 2 hop system model. For explaining the sensing error (q), we are going to use a simplified version of what was explained in Fig.4. Let us consider only the first hop of the tree in Fig.4. Having this in mind, the sensing error can be explained using Fig.5.

In Fig.5 As it was mentioned before, we are assuming a network composed by a set $\mathcal{N} = \{1, ..., N\}$ of sensors that monitor a continuous signal $x(t)$, where $t \in \mathbb{R}^+$ and $x : \mathbb{R}^+ \to \mathbb{R}$, to estimate whether a given event $\mathcal{E}$ related to $x(t)$ happens and then send this information to the network operator. Assuming that the sensors make synchronous and periodic measurements in predetermined times $t_n = n\tau$ with $n \in \mathbb{N}$ and $\tau \in \mathbb{R}^+$, we can then define a function $\theta[n]$ with $\theta : \mathbb{N} \to \{0, 1\}$ that indicates if $\mathcal{E}$ occurs at time t_n.

Due to the sensing error q, the sensors' estimation of $\theta[n]$ is going to be imperfect. For each sensor $i \in \mathcal{N}$, we have define $S_{i,j}[n]$ as the state of the system at jth level where $S_{i,j}[n] : \mathbb{N} \to \{0, 1\}$. This function represents the estimation about the given event from its individual noisy version of $x(t)$. If a sensing error at sensor i happens at t_n, then $S_{i,j}[n] \neq \theta[n]$; otherwise $S_{i,j}[n] = \theta[n]$.

After this stage that the actual state of the system has been affected by the sensing error, the sensors need to forward their binary state (whether it is the actual state of the system or it has been changed because of q) to the network operator through independent communication channels that are also subject to errors. Network operator will process the received information to determine whether $\mathcal{E}$ has occurred.

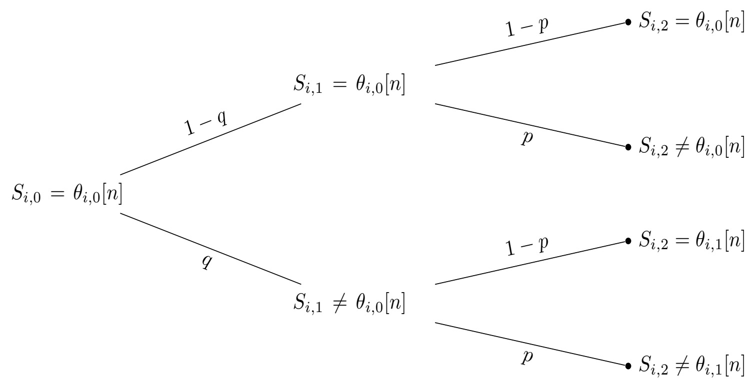

Figure 5: Schematic of the decision making tree in a one hop network with both sensing (q) and communication error (p).

Let $y_i[n]$ with $y_i : \mathbb{N} \rightarrow \{0,1\}$ be the state related to sensor i that is received by the network operator after passing through the communication link, which can be composed by only one hop or multiple hops where relay nodes forward the received information. If an error occurs in the link related to sensor i's nth measurement, then $y_i[n] \neq S_{i,j}[n]$; if not, $y_i[n] = S_{i,j}[n]$.

Network operator needs to decide whether event $\mathcal{E}$ happened at t_n based on $y_i[n]$ signals. Let $g(y_1[n], ..., y_N[n])$, with $g : \{0,1\}^N \rightarrow \{0,1\}$, denote the Boolean function that estimates the state $\theta[n]$ by the network operator so that the estimated state $\hat{\theta}[n]$ related to t_n is given by $\hat{\theta}[n] = g(y_1[n], ..., y_N[n])$. A decision error occurs whenever $\hat{\theta}[n] \neq \theta[n]$. The average error probability P_{e} of the whole process can then be calculated as:

$$P_{\mathrm{e}} = \frac{1}{\eta} \sum_{n=0}^{\eta-1} \Pr\left[\hat{\theta}[n] \neq \theta[n]\right], \tag{1}$$

where the average is taken over the different of samples such that $n \in \{0, 1, ..., \eta - 1\}$, which is related to a time window from $t_0 = 0$ and $t_{\eta-1} = T$.

The error in the final decision $\hat{\theta}[n]$ that is made by the network operator relies on the actual binary state of the system, $\theta_{i,0}[n]$, sensing and communication error and different decision functions that is made by the network operator. Different decision function studied in this chapter are explained below.

- Decision function AND: Assume that one of the binary inputs of the system is $\theta[n] = 0$, an error $\hat{\theta}[n] = 1 \neq \theta[n]$ happens only if all the inputs are one. On the other hand, if $\theta[n] = 1$ an error $\hat{\theta}[n] = 0 \neq \theta[n]$ happens if only one the inputs goes to zero.

- Decision function OR: Assume that one of the binary inputs of the system is $\theta[n] = 0$, then only one event which is equal to one is needed in order for an error to happen in the general state of the system $\hat{\theta}[n] = 1 \neq \theta[n]$. On the other hand, if the binary input of the system is $\theta[n] = 1$, an error $\hat{\theta}[n] = 0 \neq \theta[n]$ happens only if all the inputs are zero.

In the following sections, the analytical results of the average error probability of our proposed model is given based on the assumptions defined in the system model. It should be noted that first we are going to present the results based on the one hop model shown in Fig.5 and then we will expand those results for the two hop model explained in Fig.4.

According to this decision making tree in Fig.5, the sensing error q and communication error p are known, with this knowledge, different error probabilities are derived as follows :

$$\Pr[\hat{s}_{s_1} = 0 | s = 0] = (1-q)(1-p) + qp \tag{2}$$

$$\Pr[\hat{s}_{s_1} = 1 | s = 0] = (1-q)p + q(1-p) \tag{3}$$

Above equations show the probabilities of transmitting an error by smart meter 1 considering its initial state to be zero. It should be noted that $\hat{s}$ in these equations indicates the received signal. Similarly, the probabilities of transmitting an error by smart meter 1 considering its initial state to be 1 can also be derived as:

$$\Pr[\hat{s}_{s_1} = 0|s = 1] = (1 - q)(1 - p) + qp \tag{4}$$

$$\Pr[\hat{s}_{s_1} = 0|s = 1] = (1 - q)p + q(1 - p) \tag{5}$$

What we are interested to know in this research is probability of receiving an error in the final stage. Considering an one hop network with 2 smart meters ($N = 2$), the overall error probability can be defined as

$$\Pr[\hat{s} = 1|s = 0] \times Pr[s = 0] + Pr[\hat{s} = 0|s = 1] \times Pr[s = 1]. \tag{6}$$

Later on we are going to show different error probabilities according to different decision rules, due to different nature of different decisions rules studied in this research, the error probabilities related to each of the decision rules would also be different.

3.3. One hop system Model using AND decision

In this section we are presenting the analytical and simulation results related to a one hop network with 2 sensors (smart meters) using AND decision rules where sensing and communication errors are known. An illustration of the scenario studied in this part is shown in 6.

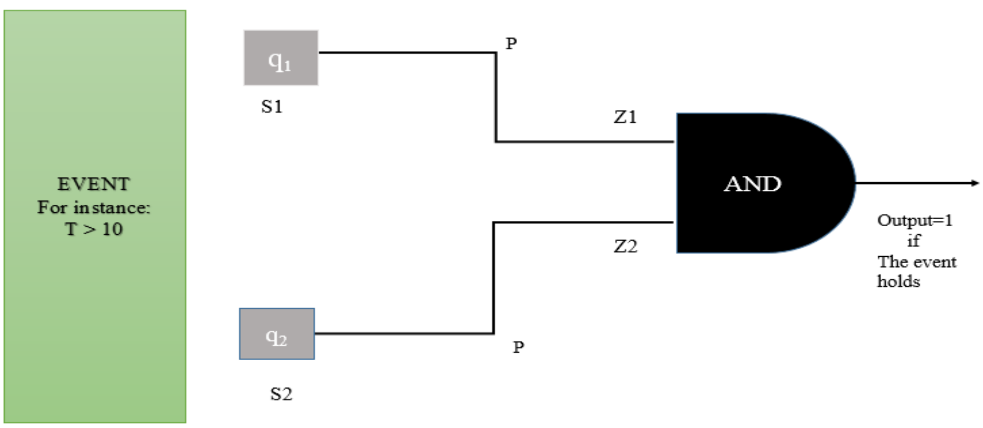

Figure 6: Illustration of the studied scenario with AND decision rule .

As mentioned before, performance of the AND gate can be described as

- **AND:** $\theta_i[n] = 0$ if at least one $\theta_{i,j}[n] = 0$ for $j \in \mathcal{N}_i$. Then, $\theta_i[n] = 1$ if all $\theta_{i,j}[n] = 1$.

Where $\theta_{i,j}[n]$ are the inputs and $\theta_i[n]$ is the aggregator i state. With these assumptions in mind, error probabilities for AND decision rules are derived as:

$$\text{Pr}[\hat{s} = 1|s = 0] = (qp)^2 \tag{7}$$

$$\text{Pr}[\hat{s} = 0|s = 1] = 1 - P[\hat{s} = 1|s = 1] = 1 - ((1-q)(1-p) + qp)^2 \tag{8}$$

Where the equation (7) shows the probability of receiving one when the input of the system is zero and equation ((8)) indicates the probability of receiving zero when the input of the system is one.

The simulation results for equations (7) and (8) are presented in Fig.7

It should be noted that in this simulation, the sensing error q is considered to be constant and equal to 20% while the communication channel error is changing form being completely reliable where $p = 0\%$ to a completely non reliable channel where $p = 100\%$.

It can be seen in this figure that when using AND decision rule, it works much better when the input of the system is zero, meaning that the average error probability of the AND decision rule is much lower when the input is zero. It is even close to zero for most of the communication error range and only when $p \geq 60\%$, even though the error starts to increase, it is still very negligible. On the other hand, when the input is 1, the average error probability is high for most of the communication error range, between 20% to 95% depending on the amount of communication error.

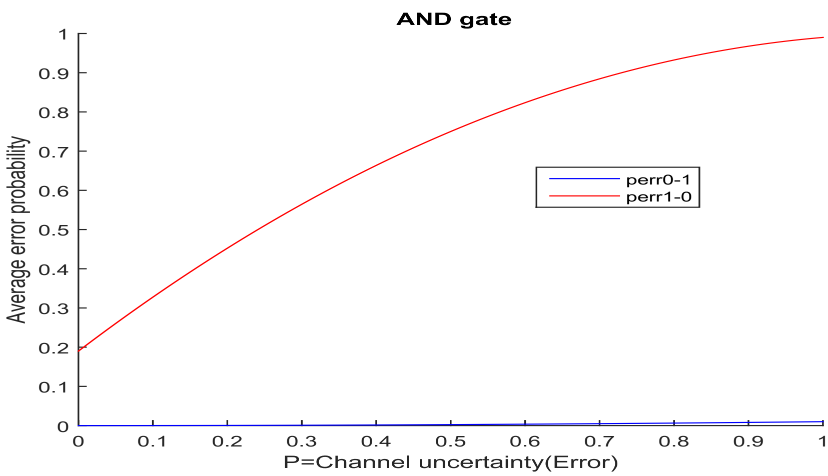

Figure 7: Average error probability of the AND decision rule for different inputs states.

This is due to the fact that AND decision function naturally favors “0" over “1". So when the input signal is “0", meaning that the above the threshold criteria is rarely met, AND decision rule works perfectly even in the presence of both the sensing and communication error. Since even if only one input would be zero, it will cause the AND gate to go to zero and no matter how high the sensing and communication errors are, they are not enouh for changing all the zero inputs to one so that AND would be 1 and therefore the average error probability is very low when we have even only 1 zero input.

On the other hand, the high average error probability of AND rule when the inputs are one is because as mentioned before, and rule naturally favors zero, so in order for it to be 1, all the inputs should indicate 1 otherwise even if only one input is zero the out put of the AND rule would be zero. It rarely happens in a system that all the inputs indicate one all together at the same time and even if they do, the presence of the sensing and communication error can cause changes in these inputs; hence, if these errors change only one of the inputs to zero, it would be enough for the output of the AND gate to be zero overall. That is why that even when $p = 0\%$ which is an ideal communication channel, the average error probability of the AND gate is still 20% and it increases dramatically as the communication error increases.

The next results that we were interested to see was the effect of increasing the number of sensor (N) in the system and how the system is going to react to this change. The results of the effect of increasing the number of sensors (users) on $P0-1$ and $P1-0$ are presented in Fig.8 and Fig.9 respectively.

It should be mentioned that in these figures, by $P0-1$ we mean the probability of receiving “1" when the input is “0" and $P1-0$ means the probability of receiving “0" when the input is one.

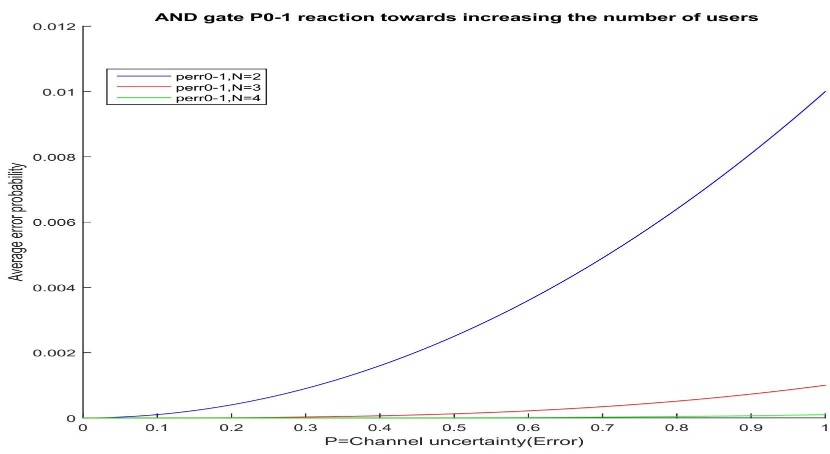

Figure 8: AND decision rule $P0-1$ reaction towards increasion increasing the number of sensors.

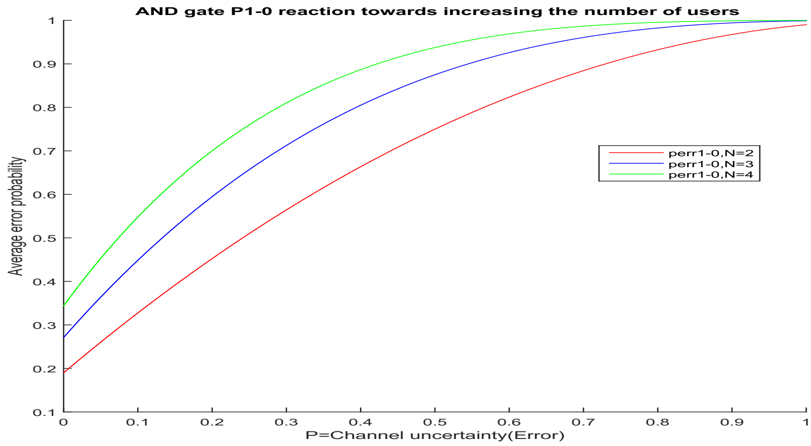

Figure 9: AND decision rule $P1-0$ reaction towards increasion increasing the number of sensors

The results obtained in this part are as expected. As can be seen in the graphs, the probability of receiving 1 when the input is zero decreased by increasing the number of sensors while the probability of receiving zero when the input is one experiences a sharp rise when the number of sensors is increasing.

The reason behind this result can be explained keeping the behavior of the AND rule in mind. When N is increasing it means that the number of inputs that might be "0" in the same time instant t_n is also increasing; hence, the probability of all the inputs being "1" at the same time instant keep decreasing and so make it almost impossible for both the sensing and communication error to change all the inputs from "0" to "1". That is why the average error probability is very close to zero when the number of users has increased to 4. Since the average error probability has already reach almost "0", we observed that further increasing the number of users will not have more effect on $P0 - 1$.

The same logic is used to explain the sharp rise in $P1 - 0$. When the number of sensors is increasing, it means that the probability of all the sensors being one at the same time which is the criteria for AND to be 1 is decreasing which will cause the probability of receiving zero when the input is one to increase highly.

In this next part, we experimented the effect of increasing the sensing error on the average error probability of the system while the communication error is changing between 0% to 100%. The results of increasing q on $P0 - 1$ and $P1 - 0$ are shown in Fig.10 and Fig.11 respectively. It should be noted that the sensing error was increased to 30%, 50% and finally 70% in these new simulations while it was fixed on 10% when running the previous simulations.

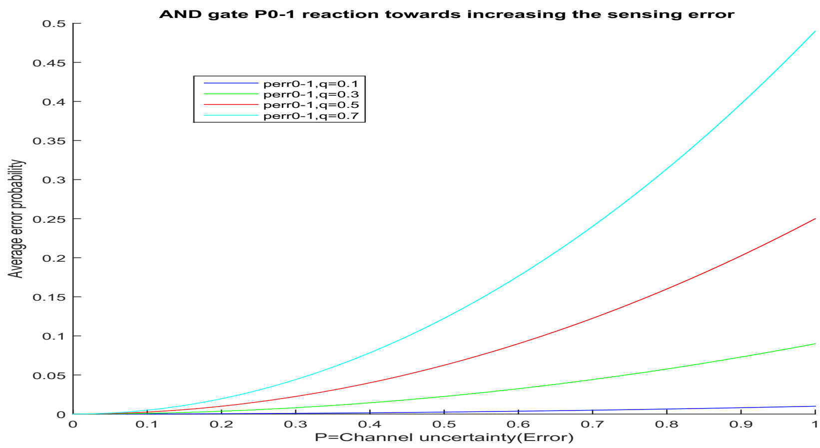

Figure 10: AND decision rule $P0-1$ reaction towards increasion increasing the sensing error.

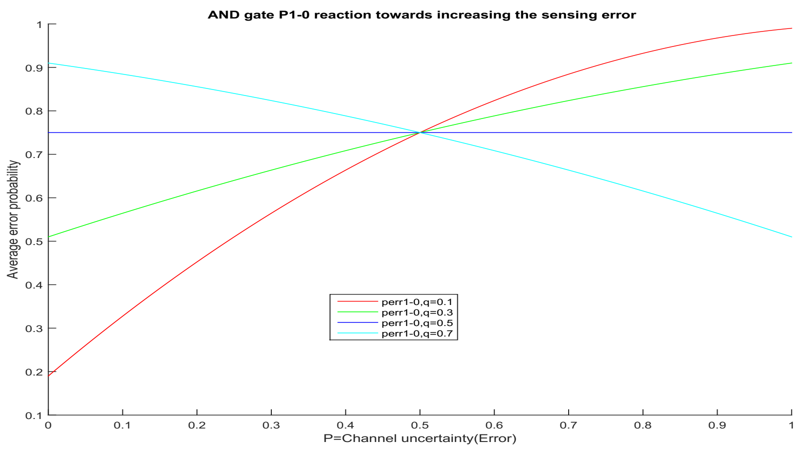

Figure 11: AND decision rule $P1-0$ reaction towards increasion increasing the sensing error.

As it was previously mentioned, the average error probability for the AND rule when the inputs are mostly zero is very low. With increasing the sensing error, it can be seen that this low error probability starts to increase specially when the communication error is more than 50%. In Fig.10, we can see this increase clearly whereas in the case where $q = 70\%$, $P0-1$ has increased to almost 50% for very high values of the communication error.

The reason seems to be that while the number of sensors is fixed, and the communication error is changing between 0% to 100%, an increase in the sensing error creates an opportunity for all the inputs to be 1 at the same time and thus increasing the probability of the out put of the AND gate to be one. This increase in the average error probability is more noticeable specifically when the communication error is high ($p \geq 50\%$). High communication error in addition to high sensing error manage to turn more and more zero inputs to one and hence increasing the average error probability when the inputs are “0".

Interesting fact that was observed about $P1-0$ was that when increasing the sensing error, after a point, the system actually starts to work better. As can be seen in Fig.11, while $0\% \leq p \leq 100\%$ for low values of q ($q = 10\%$ and $q = 30\%$) the system still works as expected where the average error probability is is increasing with an increase in sensing error, when $q = 50\%$, the average error probability becomes almost constant and equal to 75%, after this point, an increase in the sensing error will reduce the average error probability of $P1 - 0$ of the system making it decrease from 90% to 45% as the communication error is increasing. It is interesting to see that while both the communication and sensing error are very high, the system performance is actually becoming better and we obtain much lower error probability.

The reason behind this system reaction is that when the sensing error is very high (more than 50%) it manages to change several of the zero inputs to one or vice versa. On the other hand, because the communication error is also very high, it will change a lot of its received data. In many cases, this change can actually be the correction of the data it has received which had previously turned into error because of the sensing error; thus, reducing the overall average error probability of the AND gate when the inputs are mostly ones. We know that generally the average error probability of the AND function when signal is one should be high, but in this experiment we observed that when the communication and sensing errors are both high, it will result in a better system performance and lower error probability.

3.4. One hop system Model using OR decision

In this part, the analytical and simulation results for the OR gate are presented. In this part also, we are considering a simpler version of the first initial model. The model studied in this part in a one hop network with two sensors ($N = 2$). Both communication and sensing errors are known in the following simulations and analysis. An illustration of the scenario studied in this part is shown in Fig.12.

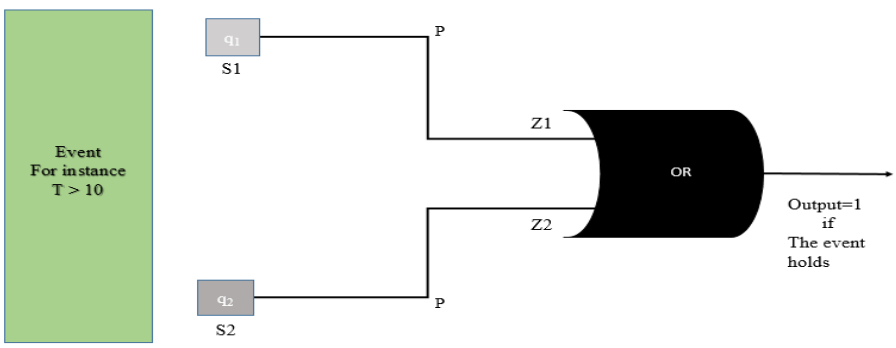

Figure 12: Illustration of the studied scenario with OR decision rule .

As it was previously mentioned, performance of the OR gate can be described as

- **OR:** $\theta_i[n] = 0$ if all $\theta_{i,j}[n] = 0$. Then $\theta_i[n] = 1$ if at least one $\theta_{i,j}[n] = 1$.

In the above expression, $\theta_{i,j}[n]$ are the inputs and $\theta_i[n]$ is the aggregator i state. With these assumptions in mind, error probabilities for the OR decision rules are derived as:

$$Pr[\hat{s} = 1|s = 0] = 1 - P[\hat{s} = 0|s = 0] = 1 - ((1 - q)(1 - p) + qp)^2 \tag{9}$$

$$Pr[\hat{s} = 0|s = 1] = (qp)^2 \tag{10}$$

It should be noted that $\hat{s}$ in these equations indicates the received signal.

In the above equations, equation (9) indicates the probability of receiving 1 when the input of the system is zero and equation (9) shows the probability of receiving zero when the input of the system is one.

Equations (9) and (10) have been simulated which the results are shown in Fig.13

It should be noted that the simulations conditions is the same as the one done in the previous part for the AND rule where the sensing error q is considered to be constant and equal to 20% while the communication channel error is changing form being completely reliable where $p = 0\%$ to a completely non reliable channel where $p = 100\%$.

As it can be seen in Fig.13 that the OR decision rules is working exactly the opposite of the AND decision rule as its performance is at its best when the input signal is one. while the AND decision rule has the lowest error probability when the input signal is zero, OR decision rule has the lowest error probability when the input signal is one.

We can see that for most of the communication error range, the average error probability of $P1 - 0$ of the OR rule is almost zero. It only increases slightly when the

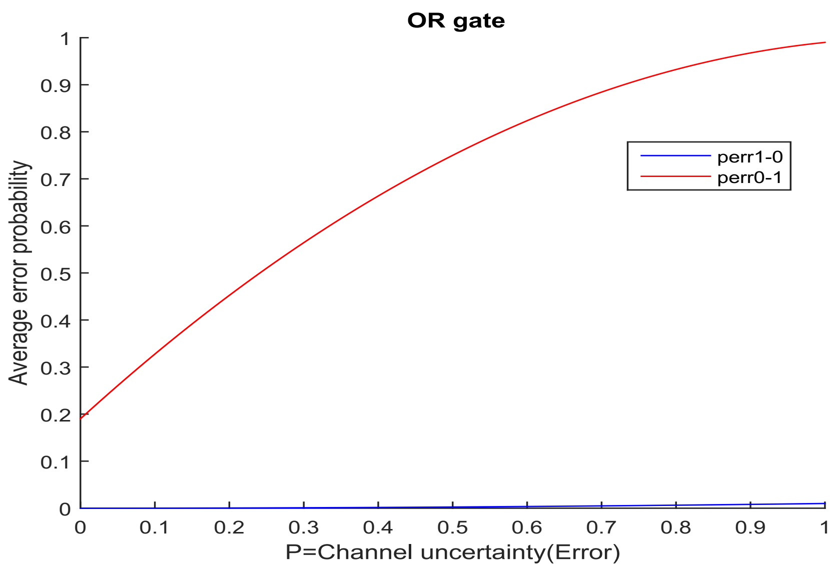

Figure 13: Average error probability of the OR decision rule for different input states .

communication error exceeds 60%, but even then, the error is still very low which was the behaviour of the $P0-1$ of the AND rule. Moreover we can see that the error rate of the $P0-1$ of the OR rule is experiencing a high error rate starting from 20% when there is no communication error and reaching almost 95% when the communication error reaches its highest point.

The reason for this kind of behaviour is that unlike the AND decision rule, the OR rule naturally favors one over zero. So we can see that in this figure, the average error probability is much lower when the signal is mostly one. We know that signal being equivalent to one means that the input signal value is above the predefined threshold that we have set for the system; hence, in a system when the signal is mostly above the predefined threshold,OR rule works very well with almost zero average error probability even in the presence of a high communication error in addition to high sensing error. In case only one input is above the predefined threshold, the overall output of the OR gate will indicate 1; thus, even very high sensing error in the first stage and very high communication error in the second stage would not have the ability to change enough ones to zero in order for the final output to be zero which will result in a very low error probability for $P1-0$ for the OR decision rule.

Furthermore, we observe a rising average error probability of $P0-1$ for OR decision rule which start from 20% when there is no communication error and the system only experience sensing error and rise up to 95% as the communication error reaches its highest values which indicates that OR rule is not a suitable choice for situations that the signal is mostly above the predefined threshold as the probability of receiving an error when the input is zero is quite high most of the times. Even very high sensing and communication errors are incapable of changing all the one inputs of the system to zeros which is the mandatory criteria for the output of the OR decision rule to be zero which is resulting in a high average error probability when the signal indicates zero at most of the time instants t_n.

In this next part, we examined the effect of increasing the number of sensors (N) on system performance and how it is going to affect the average error probability of different states of the OR decision OR. The impact of these changes on $P1-0$ and $P1-0$ are shown in Fig.14 and Fig.15 respectively.

It should be noted that in these figures as was previously mentioned, $P0-1$ means the probability of receiving 1 when the input is zero (signal below the given threshold) and $P1-0$ indicates the probability of receiving zero when the input is one (signal below the given threshold).

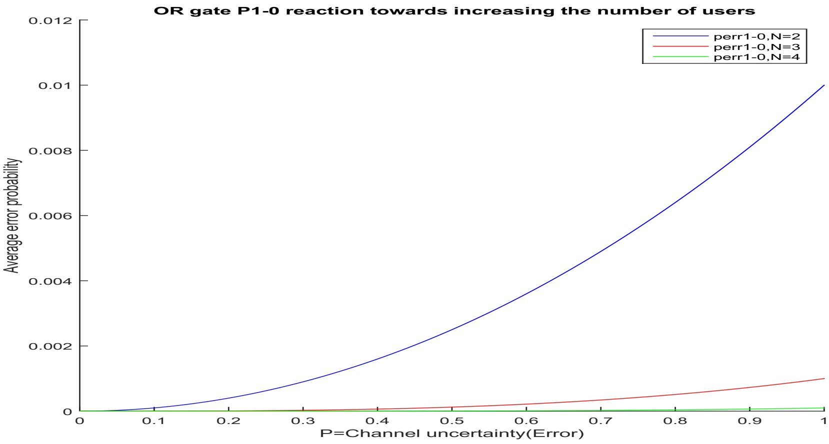

Figure 14: OR decision rule P1-0 reaction towards increasing the number of sensors.

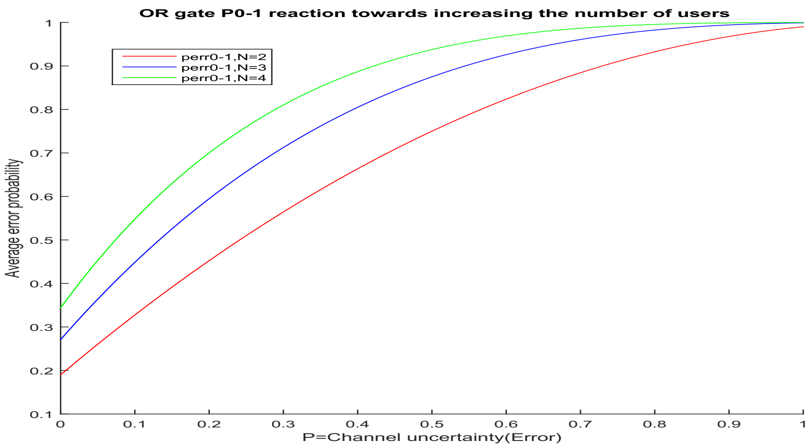

Figure 15: OR decision rule P0-1 reaction towards increasing the number of sensors.

By now we are aware of the fact that the results obtain when using the OR decision rule should be the opposite of the results of the previous part when the AND rule was used and as we can see from the above figures, the average error probability when the signal is above the threshold gets lower by increasing the number of senors while the it increases dramatically when the number of sensors get higher.

By increasing (N) we are actually increasing the probability of more input signals being above the predefined threshold at the same time instant t_n and thus increasing the number of input $1s$ at time instant t_n. Since even having only one input as one is sufficient to make the output of the OR decision rule to be one, having more one inputs means increasing the reliability of the system. That is why the average error probability for the $P1-0$ state is very low when using OR rule. We can see that $P1-0$ of the OR function has its worst performance when $N=2$ and the communication channel is absolutely unreliable and still even at its worst situation the average error probability is very low and near 1.2%. As the number of senors increases, this small error rate becomes even smaller and reaches almost zero when $N \geq 4$.

The increase in $P0-1$ can be also justified using the same logic above. As increasing the number of sensors decreases the zero inputs of the system, the probability of the output of the OR rule becoming zero will become less since the mandatory criteria for the OR rule to be zero is all its inputs to be zero at time instant t_n. As the number of zeros reduces, even high sensing and communication error in different stages would not be enough to change all the inputs into zero. Therefore, the output of the OR gate will not be zero which will result in receiving an error at the output of the network operator. As can be seen, the lowest probability of $P0-1$ for the OR rule is 20% with a perfect communication channel and 2 sensors. As the number of sensors increases, $P0-1$ also increases and when $N=4$, the system faces 35% error rate even with a perfectly reliable communication channel. At time instant t_n, for the same amount of communication error, the system is experiencing much higher error rate when the number of sensors increases.

The next set of simulation results belong to experimenting the effect of increasing the sensing error on the overall performance of the system using each of the decision rules and observing how the average error probability changes with increasing the sensing error while the communication error range is $0\% \leq p \leq 100\%$. The results of increasing q on $P1-0$ and $P0-1$ are shown in Fig.16 and Fig.17 respectively. It should be noted that the sensing error was increased to $30\%, 50\%$ and finally 70% in these new simulations while it was fixed on 10% when running the previous simulations. For better analysis of the changes in average error probability, the curves representing the error probability while $q=10\%$ are also shown in these figures.

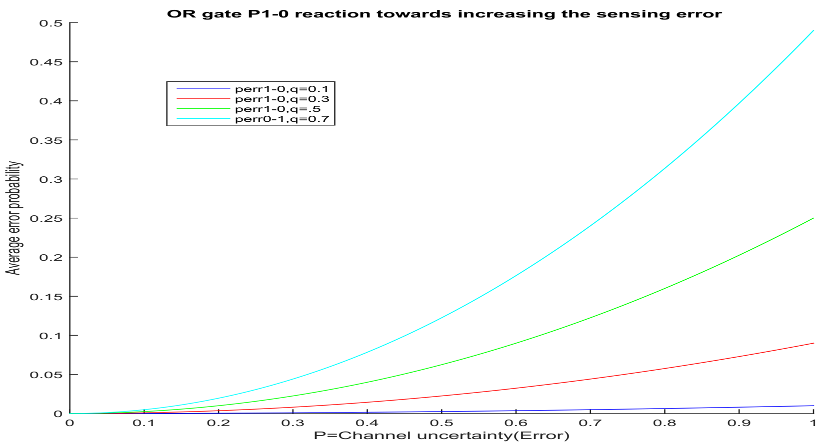

Figure 16: OR decision rule P1-0 reaction towards increasing the sensing error.

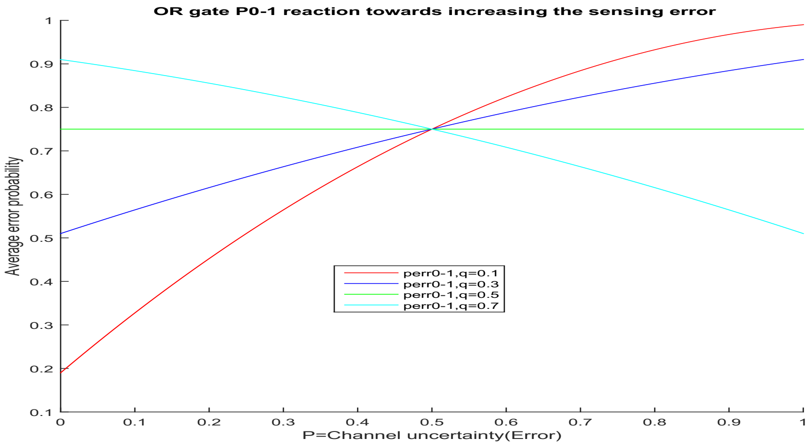

Figure 17: OR decision rule P0-1 reaction towards increasing the sensing error.

In the previous section we observed that the results of the OR decision rule is the opposite of the AND decision rule, meaning that in situations that AND rule works at its best, OR rule is working at its worst and vice-versa so that is what we are expecting to obtain in these results too.

We know from the results of the previous parts that the average error probability for the OR rule is very low when the input signal is above the predefined threshold for most of the time instants t_n. In Fig.16, we can see that by increasing the sensing error, $P1-0$ starts to increase. The average error probability becomes considerably high specially when $p \geq 50\%$ while we are increasing q to higher values. While the input signal is close to 1 most of the times, $P1-0$ is relatively low in the presence of low sensing error for the OR decision rule, but as q starts to increase, the system performance deteriorates and the average error probability for when the input signal is above the threshold reaches almost 50% when the sensing error has increased to 70% using a non reliable communication channel.

This behaviour can be justified using the same logic we used to explain the behaviour of the AND decision rule in Fig.10. In this simulation $N = 2$ and $0\% \leq p \leq 100\%$. We know that only one 1 input is enough for the out put of the OR decision rule to be one. When the sensing error is high, the probability of changing more one inputs into zero increases. Moreover, as the binary data from the sensors enter the communication channel in order to be sent to the network operator, they face a high communication error, meaning even more inputs that were not change by the sensing error are likely to change by the high communication error, therefore, increasing the probability of receiving an error when the input signal is mostly above the predefined threshold.

On the other hand, we observe a totally different behaviour in case of the $P0-1$. Although $P0-1$ which is the worst case scenario of the OR decision rule and have a high rate increases with increasing the sensing error, we can see that a very high increase in the sensing error will actually make the system functionality better when the communication error is high too ($p \geq 50\%$). In Fig.17, we observe that while the sensing error is low ($q = 10\%$ and $q = 30\%$), $P0-1$ is reacting as expected and as the communication error gets higher, $P0-1$ also increase with increasing the sensing error. When $q = 50\%$, $P0-1$ becomes almost constant (around 75%) for the whole communication error range. Further increasing the sensing error will now have a different impact on the system. In Fig.17, we can see that when the sensing error is more that 50%, the system starts to work better for higher values of the communication error. In other words, the average error probability ($P0-1$) decreases when $q \geq 50\%$ and $p \geq 50\%$. The cyan color in Fig.17 shows that the average error probability decreases from 90% to almost 45% when the $q = 70\%$ and communication channel is changing from a totally reliable channel to a completely non reliable channel. So with a very high sensing and communication error, the probability of having an error in the output of the network operator if the signal is below the predefined threshold for most the time instants t_n, becomes lower and lower.

The reason for an improved system performance in this case is that having $q \geq 50\%$ means that the sensing error has enough power to change a lot of zero inputs of the system to ones in the first stage of the network. So a lot of errors enter the communication channel as the sensors send their binary data to the network operator. In the communication channel the errors face a high communication error (when $p \geq 50\%$); hence, many of the data that have previously been change into errors by sensing error will actually get corrected by the communication error meaning that a a lot of the binary data that will reach the network operator through the communication channel will be the correct data. This phenomena will result in lower error rate of $P0-1$ for the OR rule when both p and q are very high.

We can see that the results of this part were as expected, the opposite of the results obtained when using the AND decision rule.

3.5. One hop system Model for N users

As it was explained previously, we first studied the one hop system model with only 2 users (sensors). But that is hardly the case in real life applications of smart grids. Whether we are studying HAN,NAN or WAN networks [36], there would always be more than 2 users. We presented the simulation results of the system reactions to increasing the number of sensors in the previous parts. In this part we are presenting the theoretical equations related to increasing the number of users.

we have expanded equations (7)-(10) which were obtained for a network wit 2 users for a network with N users. In equation (6) we presented the general overall error probability of the system for two users. Here we are going to expand this equation with N users for different decision rules.

3.5.1. AND decision rule average error probability with N users

In this part we are presenting the analytical result of the average error probability of the system when there are N users in the network and the network operator is using AND decision rule. The analysis in this part is also based on the one hop system model where the data encounters two kind od errors on its way to reach the network operator which are first the sensing error (q) and then the communication error (p). In these equations N indicates the number of sensors. The average error probability for the whole system is given by equation (11).

$$\mathrm{P}_e = \mathrm{Pr}[S = 0] \times (q \times p)^N + \mathrm{Pr}[S = 1] \times (1 - ((1 - q) \times (1 - p) + (q \times p)^N)) \quad (11)$$

where Pr$[S = 0]$ and Pr$[S = 1]$ represents the probability of having zero or one at the input of the network respectively. In other words, Pr$[S = 0]$ means the probability of the input signal being below the predefined threshold and Pr$[S = 1]$ means the probability of the input signal being above the predefined threshold.

When the number of users increases, ($N \rightarrow \infty$):

$$P^{AND}_{ERROR} = \mathrm{Pr}[s = 1]. \quad (12)$$

Since Pr$[S = 1]$ means the probability of having one at the input of the system,equation (12) indicates that the average error probability for the AND decision rule depends on how frequent the input signal experiences values which are above the predefined threshold.

If we want to see the individual error probabilities of the AND decision rule with N users:

$$\mathrm{Pr}[\hat{s} = 1 | s = 0] = (qp)^N \quad (13)$$

$$\mathrm{Pr}[\hat{s} = 0 | s = 1] = 1 - P[\hat{s} = 1 | s = 1] = 1 - ((1 - q)(1 - p) + qp)^N \quad (14)$$

3.5.2. OR decision rule average error probability with N users

Average error probability of the whole system when the network operator is using OR decision rule is given by equation (15) and there are N sensors in the network. In this equation too q and p indicate the sensing and communication error respectively.

$$\mathrm{P}_e = \mathrm{Pr}[S = 1] \times (q \times p)^N + \mathrm{Pr}[S = 0] \times (1 - ((1 - q) \times (1 - p) + (q \times p)^N)) \quad (15)$$

In this equation too, Pr$[S = 1]$ represents the probability of having one at the input of the system and Pr$[S = 0]$ shows the probability of having zero at the input of the network. In other words, Pr$[S = 1]$ means the probability of the input signal being above the predefined threshold and Pr$[S = 0]$ means the probability of the input signal being below the predefined threshold.

When the number of users increases, ($N \rightarrow \infty$):

$$P^{OR}_{ERROR} = \mathrm{Pr}[s = 0]. \quad (16)$$

We know that Pr$[s = 0]$ means the probability of having zero at the input of the system; therefore, equation (16) is indicating that the overall error probability for the OR decision rule is determined by how frequent the signal experiences values below the predefined threshold.

Individual error probabilities of the OR decision rule are presented below.

$$\text{Pr}[\hat{s} = 1|s = 0] = 1 - P[\hat{s} = 0|s = 0] = 1 - ((1-q)(1-p) + qp)^N \tag{17}$$

$$\text{Pr}[\hat{s} = 0|s = 1] = (qp)^N \tag{18}$$

3.6. Two hop system model analysis

It has been previously mentioned that our proposed system model is a two hop network where the information sent by the sensors first reach the relay nodes at the end of the first hop. While reaching the relay nodes, the data goes through an error-prone communication channel in a Neighborhood Area Network (NAN). Relay nodes then forward their received information to the network operator through an error-prone Wide Area Network (WAN). An example of this wo layer system model is shown in Fig.18.

In this figure, different kinds of networks for each stage of the system can clearly be seen. sensors in the houses are connected together in a home area network (HAN) [37]. When the sensors send their information from their HAN networks to their corresponding rely nodes at the end of the first hop, as it was mentioned before, the information is being send using an error-prone Neighborhood Area Network (NAN) [38].

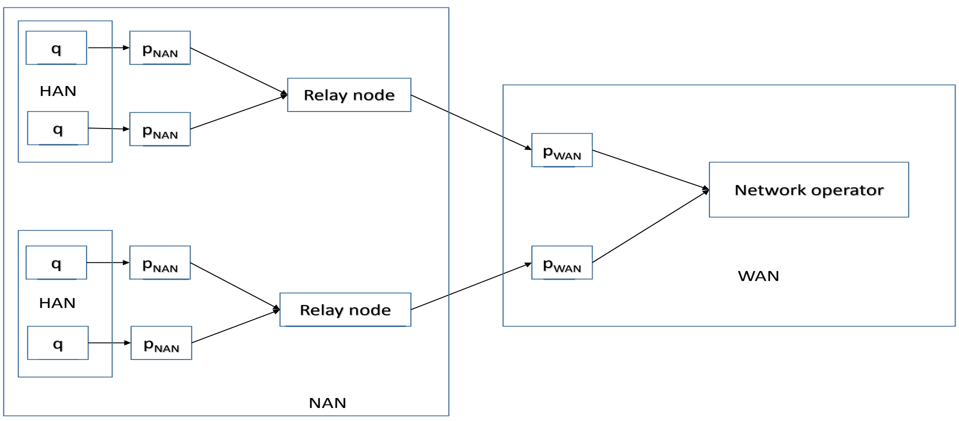

Figure 18: Two layer system model.

After this stage, it is time for the relay nodes to forward the received data to the network operator so that it could make a decision about the general state of the network. Relay nodes forward their information using an error-prone Wide Area Network (WAN) [39].

Previously in Fig.4, we saw the probability tree of a two layer system model. Keeping this decision making tree in mind, the equations for the average error probability for each of the decision rules are presented here.

3.6.1. Average error probability of a two hop system model using AND decision rule

The average error probability of the two hop system model for each different state is shown in the following equations.

$$\begin{aligned}\Pr[F=1|S=0] = {} & [((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times p_{WAN} \\ & +((1-q)\times p_{NAN}+q\times(1-p_{NAN}))^2\times(1-p_{WAN})]\times \\ & [((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times p_{WAN}+((1-q) \\ & \times p_{NAN}+q\times(1-p_{NAN})^2\times(1-p_{WAN})]\end{aligned} \tag{19}$$

$$\begin{aligned}\Pr[F=0|S=1] = 1 - {} & [[(1-(1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times p_{WAN}+ \\ & ((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times(1-p_{WAN})]\times \\ & [1-((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times \\ & p_{WAN}+((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times(1-p_{WAN})]]\end{aligned} \tag{20}$$

It should be noted that F in equations (19) and (20) means the output of the network operator which indicates the general state of the system. F depends on the input state, communication and sensing error and of course the decision making rule that is being used by the network operator. In equations (19) and (20) the average error probability of the whole system when the network operator is using AND rule is given. The simulation results based on these equations are presented

In Fig.19, the sensing error in HAN is assumed to be 20%. While data is being transmitted to the network operator, it encounters an error rate of 10% while it is going through the NAN network in the first hop of the network. In this figure we can see that WAN error is changin between 0% to 100% ($0\% \leq p_{WAN} \leq 100\%$).

As it was expected, in the two hop system model, AND rule has a low average probability when the inputs are zero and high average error probability when the inputs are one. The reason behind this has been previously explained which is due to the natural behaviour of the AND decision rule which favors zero over one.

Comparing Fig.19 with Fig.7 (one layer average error probability), we can see that although the system is working as expected, the average error probabilities obtain in the two hop model is different from the ones obtain in the one hop model. while still $P0-1 \geq P1-0$, we can see that $P0-1$ in the two hop model is much higher than the one hop model. While in the one hop model, $P0-1$ was almost close to zero for even high values of the communication error, in the two hop model, it is only close to zero for low WAN error rate and as WAN error rate increases, the system performance gets worse and $P0-1$ gets to 20% when $P_{WAN}=50\%$ and it even gets to 30% when WAN becomes completely unreliable ($P_{WAN}=100\%$).

In case of $P1-0$, it is observed that while it is still much higher than $P0-1$ when using AND rule, it also has a different behaviour compared to the one hop model. Comparing Fig.19 with Fig.7, we can see that in the two hop model, $P1-0$ has higher values for lower communication error rates than the one hop mode as $P1-0=70\%$ even when $P_{WAN}=0\%$. Although this error rate increase with increasing the WAN error, we can see that is increase is much lower than the one hop model and $P1-0$ has lower error rate for high values of communication error compared to the one hop mode. We can see that while communication error is rising from 0% to 100%, $P1-0$ only rises by 10% and reacher 80% at its highest which is lower than the one hop model.

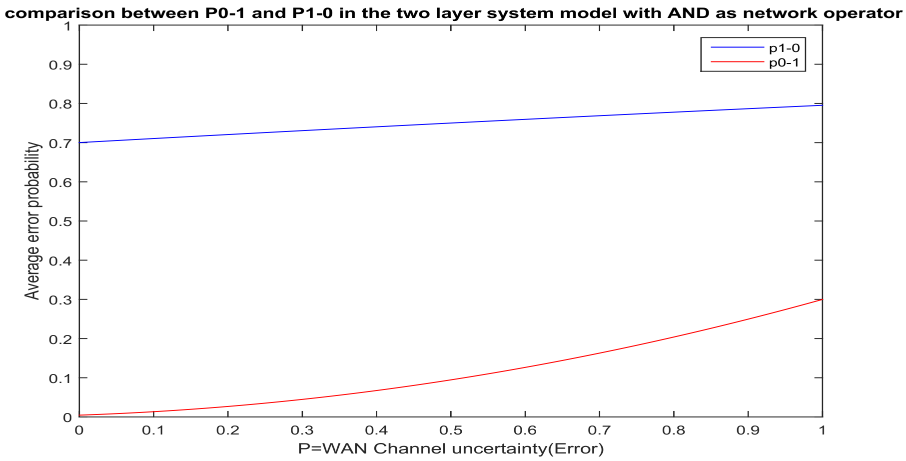

Figure 19: Average error probability of the two layer system model when AND rule is being used by the network operator while WAN error is changing.

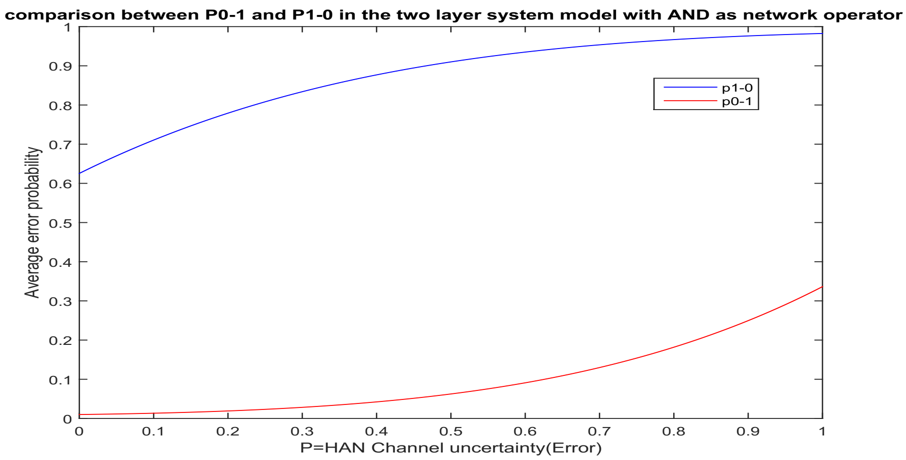

Figure 20: Average error probability of the two layer system model when AND rule is being used by the network operator while NAN error is changing.

Another simulation results of the AND rule under a different circumstance is shown in Fig.20.

In Fig.19, the sensing error in HAN is assumed to be 20%. Difference with Fig.19 is that in this scenario, WAN error is set to be 10% While $0\% \leq p_{NAN} \leq 100\%$; hence, we are studying the effect of changing the NAN error rate on the performance of the system.

We can see that the behaviour of the system complies with the theory and results that we have seen before where $P1-0 \geq P0-1$ when using AND rule. Comparing Fig.20 with Fig.7, it is shown that when NAN error in increasing, $P0-1$ bevahiour reacts almost the same way as it does for the case when WAN error was changing in Fig.19. The only difference is that in this case, for very high values of NAN error, the $P0-1$ error rate is higher compared $P0-1$ for very high values of WAN error. Moreover, while in this case, $P1-0$ has lower error rate for low values of NAN error, it experiences a sharp rise when NAN error increases to higher values and increase up to almost 100% for very high values of NAN error which indicates a 40% difference between $P1-0$ for low and high values of NAN error.

It seems that from these results, we can conclude that in the system where AND rule is being used in the network operator, the average error probability of receiving an error at the output of the network operator when the signal is below the given threshold is sensitive to changes in WAN and NAN error almost equally. On the other hand, the average error probability of receiving an error when the input signal is above the predefined threshold is more sensitive to changes in NAN error than the changes in WAN error. The reason behind this could be the fact that AND rule is much more sensitive to zeros than to ones whereas only one zero is enough to make the overall output of the AND rule to be zero, so the probability of receiving an error when the inputs are mostly zero should stay in the same range while communication error of different kind of networks changes but changes in the WAN error has less effect on the $P1-0$ of the system because it is actually the last chance that the communication error has for changing its received data into a different state and since no matter how high the WAN error is, it can not change enough data from zeros to ones so the overall output of the network operator would be one. That is why although the average error probability is still high when WAN error is increasing, the difference between the lowest and highest values of the error is less than the difference between the lowest and highest error rate of $P1-0$ when NAN error is changing. When NAN error is increasing, it changes some of the one inputs to zero or vice versa but since this data has to go through another error prone communication channel in the second hop, WAN error gets to change some of the data that has been previously changed by the NAN error and therefore increasing the average error probability of $P1-0$ when using AND rule. We can see that when the NAN error is very low, $P1-0 = 62\%$ which is less that $P1-0$ when the WAN error is changing, but as NAN error gets higher, $P1-0$ increases to higher values compare to $P1-0$ error rate when WAN error is changing.

3.6.2. Average error probability of a two hop system model using OR decision rule

The average error probability of the two hop system model when OR rule is used by the operator is shown in the following equations for each different states.

$$\begin{aligned}\Pr[F=0|S=1] = {} & [((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times p_{WAN} \\ & +((1-q)\times p_{NAN}+q\times(1-p_{NAN}))^2\times(1-p_{WAN})] \\ & \times[((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times \\ & p_{WAN}+((1-q)\times p_{NAN}+q\times(1-p_{NAN})^2\times(1-p_{WAN})]\end{aligned} \tag{21}$$

$$\begin{aligned}\Pr[F=1|S=0] = 1-{} & [[(1-(1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times p_{WAN}+ \\ & ((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times(1-p_{WAN})]\times \\ & [1-((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times \\ & p_{WAN}+((1-q)\times(1-p_{NAN})+q\times p_{NAN})^2\times(1-p_{WAN})]]\end{aligned} \tag{22}$$

As it was expected, the equations obtained in this part are the opposite of the equations obtained in the previous part. Here like before, F represents the general state of the network at the output of the network operator which is determined by the sensing error (q), NAN and WAN error and most importantly the decision rule that is chosen by the network operator.

Using equations (21) and (22), we are able to calculate the average error probability of the network when the network operator is using OR decision rule. The simulation results presented in Fig.21 are the results of simulating the network using the above equations while WAN error is changing.

Like in Fig.19, in this figure too $q = 20\%$ in the HAN network. After some of the data is changed because of q in the sensors in HAN, it goes through the error

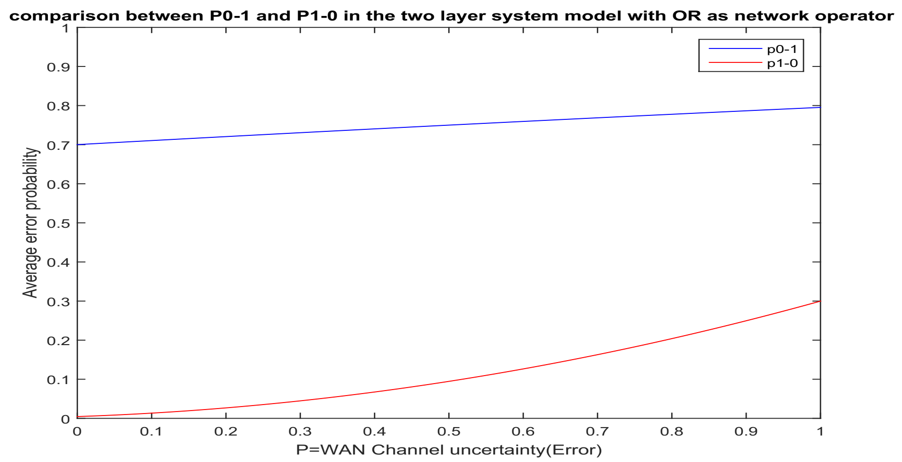

Figure 21: Average error probability of the two layer system model when OR rule is being used by the network operator while WAN error is changing.

prone NAN networks where it faces a communication error equal to 10%. After this encounter, data has reached the relay nodes which is the end point of the first hop. After this point, the information which now has been tempered with by the communication error in the NAN network goes through the next hop which is an error prone WAN network where error is $0\% \leq p_{WAN} \leq 100\%$.

It is shown as expected for the OR decision rule, $P0-1 \geq P1-0$. The average error probability for the two hop system model when the signal is mostly above the predefined threshold when OR decision rule is being used by the network operator is much lower than the average error probability when the input signal is mostly below the given threshold of the system. This is due to the fact that OR decision rule is much more sensitive to one than to zero.

If we take a look at Fig.14, we can see that the average error probability of the system when the signal is above the predefined threshold is almost zero for most of the NAN communication range, while in Fig.21 still $P1-0 \leq P0-1$ holds, the error rate of the $P1-0$ is much higher compared to $P1-0$ in the one hop model. In Fig.21, $P1-0$ is only close to zero for very low values of the WAN error. As the WAN error increases, we see that $P1-0$ increases to 30% when the WAN error is at its highest value.

We observe that while the performance of the system when the input signal is mostly below the predefined threshold of the system, is much worse than the performance of the system when the input signal is mostly above the predefined threshold, the error rate of $P0-1$ is different in this two hop model is different from the error rate of $P0-1$ in the one hop model. $P1-0 = 70\%$ in this two hop model even when there is no WAN communication error in the system which is much higher than the error rate for the same amount of communication error in the one hop model. Even though the error rate is higher for low WAN error rate in the two hop model, $P1-0$ is lower for the higher WAN error values and it increase by only 10% and reaches 80% at the worst case of the WAN error rate.

The next simulation results presents the behaviour of the two hop network while NAN error is changing.

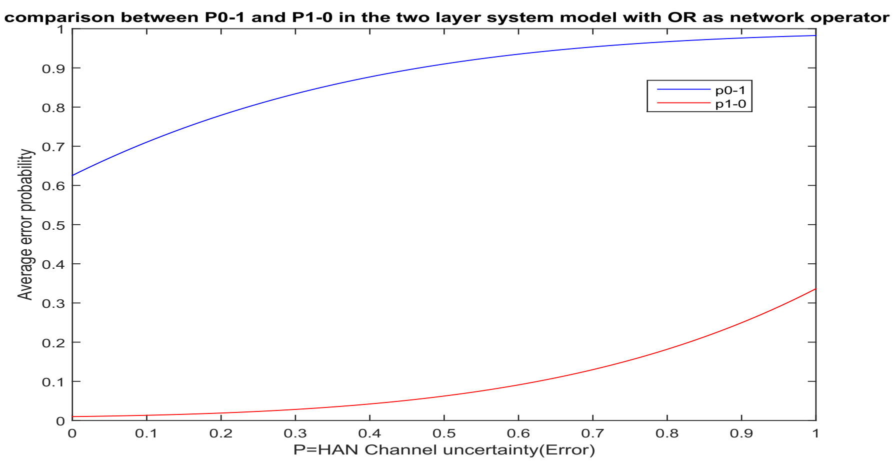

Figure 22: Average error probability of the two layer system model when OR rule is being used by the network operator while NAN error is changing.

In Fig.22 the situation is reversed compare to Fig.21. In this part, sensing error which is in the HAN network is again set to be 20%. In this scenario, WAN error is also fixed and is 10%. Since we want to study the effect of changes of the NAN error on the system, it is changing between 0% to 100% ($0\% \leq p_{NAN} \leq 100\%$).

When NAN error is changing, system is still behaving as we saw before for the OR function where $P0-1 \geq P1-0$. The behaviour of the average error probability of the system when the input signal is above the predefined threshold for most of the time instants t_n is almost the same as for the case that the WAN error is changing when using OR rule. The difference is that when the NAN error rate is very high, $P1-0$ is slightly higher than $P1-0$ when the WAN error was changing (By 5%). On the other hand, $P0-1$ has a different reaction to increasing the NAN error rate compared to its reaction to changes in the WAN error. $P0-1$ in Fig.22 has a lower value for low values of NAN error (around 62%). Although like what we observed in Fig.21, $P0-1$ is increasing when NAN error is increases, contrary to the same Fig.21 this rise is high and it increases the rate of $P0-1$ by almost 40%.

From the results presented in this part, we observe that when using the OR decision by the network operator, the probability of receiving an error at the output of the network operator when the input signal is below the predefined threshold for most the time instants t_n is much more sensitive to changes is the NAN error rate than it is to changes in the WAN error rate since he rise in $P0-1$ is much more higher when NAN error is changing compared to the scenario that WAN error is changing. We can see that when using the OR decision rule, while $P1-0$ increases with increasing both NAN and WAN error range, there is not much of a difference in the range that his rise happens since it increases from 0% to 35% when NAN error is increasing and 0% to 30% when WAN error is increasing. This reaction too can be justified keeping the natural behaviour of the OR rule in mind. Since it favors ones over zero and only having one 1 input is enough for the overall output of the network operator to be one, when the signal is above the threshold, although the system is sensitive to increases in different errors, it is not very sensitive to each individual error and it reaction to increasing both of the network error rates stays almost the same when the input signal is above the threshold. On the other hand, we observe that $P0-1$ when using OR rule is much more sensitive to changes in NAN error rate compare to changes in WAS error rate. As it was previously explained, When NAN error is fixed, it does not have the power to change a lot of zero inputs to one which is the sufficient criteria for the whole network to be zero, so when the information goes through the WAN network, while WAN error manages to changes some of its received information specially at its highest values, it can not effect the average error probability a huge amount and that is why we only see a 10% increase in $P0-1$ when WAN error is changing but when NAN error is increasing and is not fixed anymore, it changes some of the inputs to errors and then send these information to the WAN network when they encounter another error which is high enough for increasing the error rate but not high enough for correcting the previously changed errors and thus reducing the error rate. That is why we see a sharp increase (by almost 40%) in $P0-1$ rate when NAN error is changing compared to when WAN error is changing.

In the next chapter, we will see the simulation results of implementing this proposed system model on actual real time data.

4. PRACTICAL IMPLEMENTATION

The power grids that are being used nowadays have been active for more than a century now. Since they were designed a long time ago, they were only responsible for transferring electricity from the utilities to the costumers [40]. With advancement of technology, these traditional power grids do not meet the requirements needed for our daily lives anymore and this is causing many problems such as more grid failures, voltage sags and blackouts [41]. That is why the concept of smart grids have been introduced. Empowered by the advances of information and communication technologies, it has the ability to mitigate the current problems that the traditional power grids are facing [4], [42].

Integration of the communication networks with the traditional power grids is what making the traditional power grids into smart girds since it makes the communication between different grid elements possible [43], making the communication network one of the most important part of a smart grid [22].

That is the reason that in this chapter we are focusing on a new way of implementing the communication network in a smart grid.

4.1. System Model

In the previous chapter, we analyzed our proposed system model from a theoretical point of view. First we started by analyzing the one hop system model in order to be able to get a general idea about the behaviour of different decision rules under different circumstances in the system. The average error probability equations related to each decision rules were derived and simulated to see the effect of each system parameter on the average error probability of the network. Then we moved on to study the two hop system model where the information is send from sensors to aggregators and from aggtregators to the network operator. It was mentioned that the aggregators studied in the previous chapter only work as relay nodes and they only forward the information they receive to the network operator in the next hop.

In this chapter, individual smart-meters periodically monitor the average power demand of their respective households to inform the system operator if it is above a predetermined level using only a 1-bit signal. The communication link connecting these smart meters together is, like the previous chapter, error prone. The communication network in this model is consist of two layers. First layer is composed by individual smart meters which are connected to their corresponding aggregator. Second layer is defined as different aggregators connected to the system operator. This distribution system operator needs to estimate whether the average power demand in a given period is above a predetermined threshold using an 1-bit memoryless scheme. At the end point of each stage, a decision is made about the aggregated state of the system. This decision is based on hard decision rules defined by different decision rules. In this chapter, AND, OR and also MAJORITY decision rules are studied.

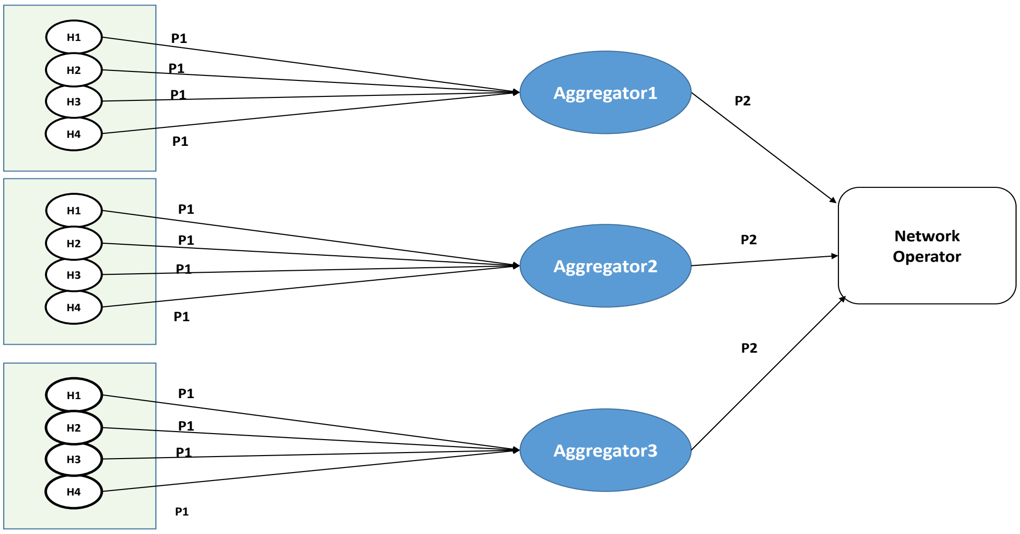

Figure 23: Example of the scenario studied in this chapter.

It should be noted that the difference between the model studied in this chapter and the one that was studied in the previous chapter is that in the previous model, there was no decision making happening in the aggregators as they were only working as rely node, but in this model, the decision making is actually happening in two stages of the networrk. First, at the end of the first hop by the aggregators, the second decision is then made by the network operator at the end of the second hop about the general state of the system.

As it was mentioned earlier, here we focus on analyzing a non-critical application where the system operator needs to estimate whether the average power demand in a given period (e.g. 15 minutes) of the distribution grid is above a predetermined threshold. Our goal is to build an efficient communication system with simple and low cost implementation. To do so, we follow our previous work [44] (which was also explained in the chapter 3) to build a WSN in two hops such that individual smart meters send to their respective aggregator an 1-bit message indicate whether the individual average power demand is above a given threshold. The aggregators then decide about their state based on the received information and send their decision as an 1-bit signal to the system operator, which its turn decides if the power demand is above the threshold in the same way.

The reason we are choosing to use wireless sensor networks (WSN) [45] for implementing our idea is that wireless sensor networks which build the core of IoT, are becoming widespread in almost any possible application area ranging from energy systems to sleep monitoring [46], [12].

The system model studied in this chapter is shown in Fig.23.

In this figure, Smart meters monitor the average power demand in order to determine its binary state $\theta_{i,j}[n]$ at time t_n. The meters are associated with 3 aggregators that decide their state $\theta_i[n]$ based on the inputs from 4 household data. The aggregators then send their state $\theta_i[n]$ to the system operator that will decide about the global state $\theta[n]$. , communication error of different hops are presented in this figure as $P1$ and $P2$

where $P1$ represents the communication error in the NAN network and $P2$ represents the communication error in the WAN network

First assumption to be considered in this model is that in a given period of time, the consumption is above a predefined threshold with a given probability. Moreover, the channel is modeled as a binary symmetric channel (BSC). Based on the mentioned assumptions, the error probability of the system operator is derived. What we are trying to see here is how efficient is our proposed one bit signaling model under different circumstances. We are looking for the best strategy given the channel error probabilities is determined by how frequent the meters experience the consumption above the defined threshold. The impact of increasing or decreasing the defined thresholds and also the communication error and the way the decision rules react to these changes are also studied in this chapter.

As can be seen, this model consists of two layers, first layer is the connection between the houses and network aggregators. There are a total of 12 houses studied in this chapter. We have grouped them in 3 different networks consisting of 4 houses each. In this chapter, we choose to build a generic communication model for non-critical applications in the distribution level of the power grid. In this way, our approach does not focus on high reliability or low latency, but rather on a cheap way to estimate the average power demand without harming the communication network with huge amounts of data (e.g. [33]).

We consider a case where the smart-meters inform the aggregator whether their average power demand in predetermined time periods is above or below a given threshold. Aggregators process the received information using AND, OR and MAJORITY (memoryless) logical operations and send the processed information to the the system operator. The system operator then decides based on the same logical operation if the aggregate average power demand is above or below the threshold. We test our approach using 12 daily demand profiles taken from the database "The Reference Energy Disaggregation Data Set" (REDD) [47]. Thus, as depicted in Fig.23, every 4 houses connect to one aggregator, therefore we assume 3 aggregators besides the network operator. Information obtained from this study can then be used in, for example, cloud computing applications [21] such as forecasting, demand-side management, peer-to-peer energy trading.

Let us assume a network composed by a set $\mathcal{N}_i = \{1, ..., N_i\}$ of smart meters of a given group of consumers (*prosumers*) i composed by N_i elements, which are associated with aggregator $i \in \mathcal{N}$ where $\mathcal{N} = \{1, ..., K\}$ is the set of aggregators. Each meter $j \in \mathcal{N}_i$ needs to inform aggregator i in predetermined times $t_n = t + n\tau$ if its individual average power demand, $P_{i,j}(t_n)$, is above or below a given threshold γ. Let $\theta_{i,j}[n]$ be the function that indicates whether $P_{i,j}(t_n) > \gamma$.

Note that since the communication links connecting the smart meters in different houses to their corresponding network aggregator are defined as a neighborhood area network (NAN) and the links connecting the network aggregtors to the system operators are defined as wide area networks (WAN), in our analysis, we have assumed that the errors associated with each of these networks are almost the same, so the error probability in the communication links connecting the houses to the aggregators are considered almost the same in all the three NAN networks (p_1). Also,since the links connecting the aggregators to the network operator are all part of the same WAN networks, the links error are considered the same (p_2).

We assume that smart-meter $j \in \mathcal{N}_i$ sends its state $\theta_{i,j}[n]$ to aggregator i through a binary symmetric channel (BSC) [35, Ch.7] with error probability p_1 (the subscript "1" indicates the first communication hop). Based on such information, aggregator i decides its state $\theta_i[n]$ using hard-decision rules AND, OR or MAJORITY from the inputs $\theta_{i,j}[n]$. Aggregators $i \in \mathcal{N}$ then needs to send its state $\theta_i[n]$ to the system operator in a binary symmetric channel with error probability p_2. With the information from all aggregators in hand, the operator similarly proceeds to decide the global state $\theta[n]$ based on AND, OR or MAJORITY logic operations. As shown in Fig.23,we assumed a network of houses which are connected together in a star like network topology. The aggregators in Fig.23 are the first stage that different decision rules are implemented and the network operator is the second place that the decision making will happen, unlike the model studied in the previous chapter where decision making only happened in the network operators.

Our goal in this chapter is to understand how $\theta[n]$ is affected by the decision rules (AND, OR or MAJORITY), channel error probabilities p_1 and p_2, input state , the number N_i of smart meters for aggregator i and the number of aggregators K. Moreover,the effect of the predefined threshold of the average power demand on the error probability of the received signal and the behavior of each of the decision rules is also studied .

As it was mentioned above,three decision rules are studied in this chapter which can be explained as:

- **AND:** $\theta_i[n] = 0$ if at least one $\theta_{i,j}[n] = 0$ for $j \in \mathcal{N}_i$. Then, $\theta_i[n] = 1$ if all $\theta_{i,j}[n] = 1$.
- **OR:** $\theta_i[n] = 0$ if all $\theta_{i,j}[n] = 0$. Then $\theta_i[n] = 1$ if at least one $\theta_{i,j}[n] = 1$.
- **MAJORITY:** As aggregator i has N_i inputs, then $\theta_i[n] = 0$ if $\sum\limits_{j \in \mathcal{N}_i} \theta_{i,j}[n] = 0 < N_i/2$ and $\theta_i[n] = 1$ if $\sum\limits_{j \in \mathcal{N}_i} \theta_{i,j}[n] = 0 > N_i/2$. If $\sum\limits_{j \in \mathcal{N}_i} \theta_{i,j}[n] = 0 = N_i/2$, then $\theta_i[n]$ is randomly selected with 50% of chance.

In the results that will b presented later we will show how sensitive the average error probability is different communication errors and also that the decision rule that decreases the error probability is determined by how frequent the meters experience the consumption above the threshold in addition to the predefined threshold.

For better understanding of the communication error, let us consider $S_{i,0}[n]$ to be the signal sent by sensor i and $S_{i,j}[n]$ be the state of the system at jth level. j is defined as the number of hops in the network and can be $j = 1, ...M$, the studied model in this chapter was said to consists of two hops, one from the networks of smart meters to the network aggregators and the other one from the networks aggregators to the network operators. At every hop,the binary state of the network ($\theta_{i,j}[n]$) is forwarded to the next hop using the communication channels which as have been mentioned before, are subjected to error.

As we know, in this model we assume the communication channel to be binary symmetric. Error probability p_j is what causes the output of this channel to be different from the input. The assumption is that the errors happening in different layers of this model at different time-steps t_n are independent from each other; Hence, the state of the system will be transmitted to the next well with error probability p_j where j defines the number of the layer (hop) of the system. This means that the state $S_{i,j} = S_{i,j-1}$ with error probability $1 - p_j$ and $S_{i,j} \neq S_{i,j-1}$ with error probability p_j. It should be noted that unlike the previous model, in this model there in no sensing error q and the only error in this model are the different layers communication errors.

The error in the final decision $\hat{\theta}[n]$ that is made by the network operator relies on the actual binary state of the system, $\theta_{i,0}[n]$, communication errors and different decision functions that have been made through out the networks in each hop.

4.2. Average error probability

As it was previously mentioned, in the first layer, each network has a predefined threshold, if the average power demand of each house would be above this threshold, that house would send 1 to its corresponding aggregator. In other words, each house would send 1 if $P_{i,j}(t_n) > \gamma$. where γ is the predefined threshold which is the same as the NAN network threshold.

Moreover, there is also another predefined threshold which is used by the network operator in the last hop in order to calculate the average error probability of the whole network. The same procedure explained before for deciding the state of the network is also used here; Hence, in this hop it is defined as $\bar{P}_{i,j}(t_n) > \sum\limits_{i=1}^{K} N_i \times \mathrm{P}_{th}$. In this equation, $\sum\limits_{i=1}^{K} N_i * \times \mathrm{P}_{th}$ is the predefined threshold of the network operator. N here indicates that the threshold of the second hop is a multiple of the thresholds of each network of houses in the first hop. Based on the explained expressions, the general problem is formulated as follows.

$$\begin{cases} \theta = 1, & \text{if} \sum_{i=1}^{K} \sum_{j=1}^{N_i} P_{i,j}(t_n) > \sum\limits_{i=1}^{K} N_i \times \mathrm{P}_{th} \\ \theta = 0, & \text{otherwise} \end{cases} \tag{23}$$

For better understanding of how the error is identified in this system, Let $s[n]$ be the binary function denoting whether $\sum\limits_{i \in \mathcal{N}} \sum\limits_{j \in \mathcal{N}_i} \bar{P}_{i,j}(t_n) > P_{\text{th}}$. In this case the value of $s[n]$ indicates the actual state of the network at time t_n and, therefore, this shall be used as the basis of comparison for the communication scheme proposed in Section. By doing

Table 2: Example of how to compute the average error probability

n	$\sum \bar{P}_{i,j}(t_n)$	P_{th}	$s[n]$	$\theta[n]$	Error
1	3956	7500	0	0	No
2	7843	7500	1	1	No
3	11373	7500	1	0	Yes
4	7005	7500	1	1	No
5	7897	7500	1	1	No
6	6353	7500	0	1	Yes

so, we can define an error event associated with the measurement done at t_n whenever $\theta[n] \neq s[n]$. An example of how the average error probability is computed can be seen in table 2.

As previously discussed, $\theta[n]$ is built to be a simple and cheap estimation of $s[n]$, which error events would still happen even with perfect communication channels. Including errors in the communication will further increase the uncertainty of the estimation $\theta[n]$. Herein, we are interested on the average error probability over n such that

$$P_{\text{er}} = \frac{1}{n_{\text{max}}} \sum_{n=1}^{n_{\text{max}}} \mathbb{1}[\theta[n] \neq s[n]], \tag{24}$$

where $\mathbb{1}[\cdot]$ is the indicator function and n_{max} is the number of measurements considered.

Table 2 is an example of our framework by showing the average power demand of 12 households considering 6 measurements such as number of the house, average power demand, threshold, etc. The state $s[n]$ indicates if $\sum_{i,j} \bar{P}_{i,j}(t_n) > P_{\text{th}}$, while $\theta[n]$ is the estimation considering the proposed 1-bit signaling including communication errors. In this example, $n_{\text{max}} = 6$ and $\sum_{n=1}^{n_{\text{max}}} \mathbb{1}[\theta[n] \neq s[n]] = 2$ (i.e. two error events happened). Then, the average error probability is $P_{\text{er}} = 2/6 = 33.3\%$.

4.3. Numerical results

In this part we are implementing our proposed model to compute the average error probability which was presented in the previous part to examine the error probability of each of the decision rules AND, OR and MAJORITY when they are being used in this model. Moreover, the impact of communication error in both of the NAN and WAN networks are also studied. As it was previously mentioned, for the analysis that has been carried out in this chapter, we use "The Reference Energy Disaggregation Data Set" (REDD) [47] database to generate a 15-minute average power demand over a timespan of 24 hours (one day) for 12 different daily profiles, yielding $n_{\text{max}} = 96$.

In order to be able to carry out our analysis on the data we obtained from REDD, the algorithm below was used. Since the actual code is very long, this algorithm is a summerized version of the actual code.

Algorithm 1 Average error probability calculation algortihm

```
procedure MyProcedure
    for nan.thre:=100 to 1000 step 50 do
        Sheets("sheet(i)").Range("B", i).Value ← hous.consumption(j)
        total.consumption = house.consumption(1)... + house.consumption(12)
        if total.consumpion > total.threshold then
            real.output = 1
        else
            real.output = 0
            EndIf;
        if house.consumpion(i.j) > threshold(j) then
            Binary.consumption(i) = 1
        else
            Binary.consumption(i) = 0
            EndIf;
        for j:=1 to 1000 do
            Applying first hop communication error
            if Rnd() ≤ nan.communication.error(j) then
                Binary.consumption.err(i) = 1 − Binary_consumption(i)
            else
                Binary_consumption_err(i) = Binary_consumption(i)
                EndIf;
        Applying first hop OR/AND/MAJORITY function
        Applying second hop communication error
        if Rnd() ≤ wan.communication.error(j) then
            OR/AND/MAJ.fnc.ly1.2(i) = 1 − OR/AND/MAJ.fnc.ly1(i)
        else
            OR/AND/MAJ.fnc.ly1.2(i) = OR/AND/MAJ.fnc.ly1(i)
            EndIf;
        applying second hop OR/AND/MAJORITY function
        if OR/AND/MAJ.fnc.ly2 ≠ real.output then
            count.error.OR/AND/MAJ = count.error.OR/AND/MAJ + 1
            EndIf;
        Next j.
        total.error.AND/OR/MAJ.per.input = total.error.AND/OR/MAJ.per.input + count.error.OR/AND/MAJ/1000
        Sheets("sheet3").Range("AH", index.nan.thre).Value = total.error.and.per.input/95
        Next nan.thre.
```

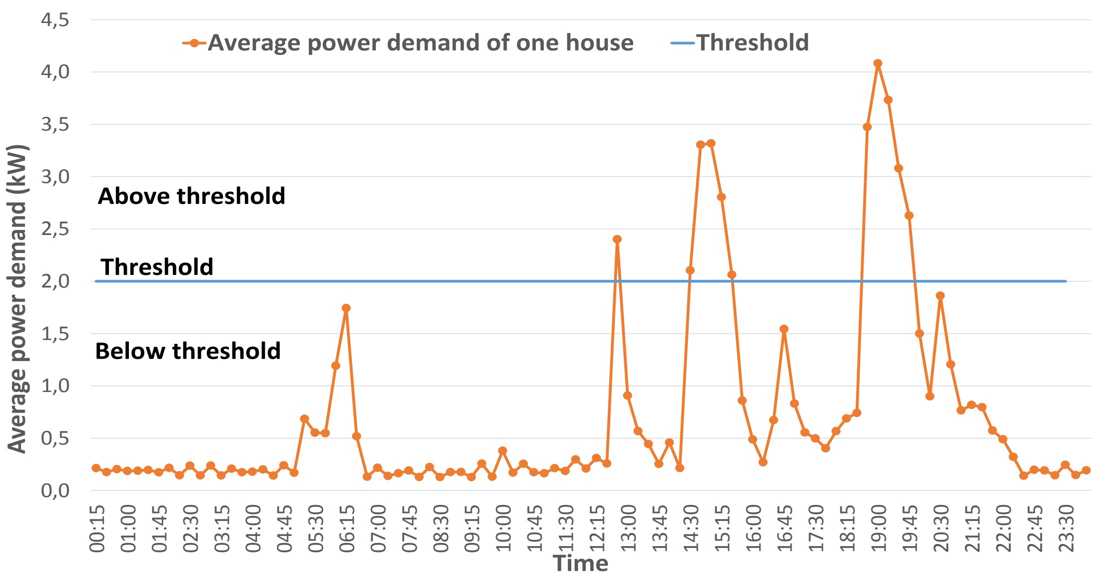

Figure 24: An example of the average power demand for one house $P_{i,j}(t_n)$ in the network and its corresponding threshold γ. If $P_{i,j}(t_n) > \gamma$, then $\theta_{i,j}[n] = 1$; otherwise, $\theta_{i,j}[n] = 0$.

As it has been mentioned earlier, each of the three networks of houses in the first hop has a predefined threshold and if the average power demand of each house in one of these would be above this threshold, that house would send 1 to its corresponding aggregator. Fig.24 shows how the individual smart meter codes its average power demand $P_{i,j}(t_n)$ into $\theta_{i,j}[n]$. If the individual demand $P_{i,j}(t_n)$ is above the individual threshold γ, then $\theta_{i,j}[n] = 1$; otherwise, $\theta_{i,j}[n] = 0$. We proceed similarly with all 12 households to obtain the states $\theta_{i,j}[n]$ that are the communication system inputs as described earlier.

There are 3 network of houses in this study. Fig.24 shows the average power demand of one of the houses in one od these networks. It can be seen in this figure that the predefined threshold which is also the same as NAN threshold here is set to be 2kW. This threshold is determined based on the average power demand of each different houses in one network and therefore it is different in each of the studied network. So this threshold is the same for the houses of one network but it is different compared to the other networks.

Also, it has been explained that there is also a predefined threshold in the second hop which is used by the network operator for calculating the average error probability of the whole network and deciding on the general state of the whole system. Fig.25 illustrates this threshold which is the same as WAN threshold. Fig.25 also exemplifies how the state function $s[n]$ which was explained in equation (24) is obtained. The aggregated average power demand curve is plotted and compared with the system operator threshold P_{th}. If the sample at time t_n is greater than P_{th}, then $s[n] = 1$; otherwise $s[n] = 0$. As discussed before, $s[n]$ provide the actual system state that the estimation $\theta[n]$ shall be compared.

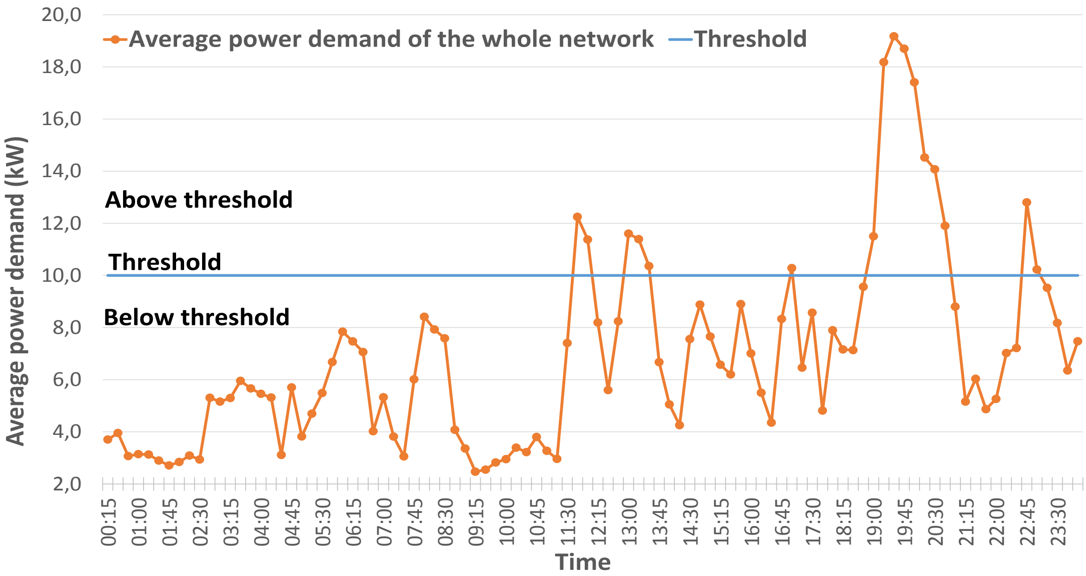

Figure 25: An example of the aggregated average power demand $\sum P_{i,j}(t_n)$ and its corresponding threshold P_{th}. If $\sum P_{i,j}(t_n) > P_{\text{th}}$, then $s[n] = 1$; otherwise, $s[n] = 0$.

In this figure, it can be seen that the threshold of the network operator is set to be 10kW. This threshold was selected according to the aggregated average power demand ($\sum P_{i,j}(t_n)$) of the whole network (all 12 houses).

The simulations were ran for different scenarios in this part; thus, several results were obtained regarding each scenario, here we are presenting the results which are the most interesting ones from a designer point of view.

First, let us focus on the effects of the individual threshold γ and the average error probability P_{er}, which is given by equation (24). The effect of changes in the NAN threshold on the average error probability of the whole system in an unreliable network are presented in Figs.26-28. Note that in the following results, the decision making rules used in both of the network hops are the same. Figs present P_{er} as a function of the individual threshold γ for communication error probabilities $p_1 = 20\%$ (NAN communication error) and $p_2 = 10\%$ (WAN communication error), and the system operator thresholds $P_{\text{th}} = 5, 7.5$ and 12.5 kW, respectively. We consider the threshold γ ranges from 0.1 to 1 kW.

Fig.26 represents a scenario were the threshold of the WAN network is set too low. So $WAN_{thre} = 5000KW$ (WAN threshold) and the NAN network threshold is changing between 100 to 1000 KW.

It should be mentioned that the following results were obtain by running a Monte Carlo simulation for each of the scenarios. (1000 snapshots). In Fig.27 and 28, NAN threshold was increased to 7500 and 12500 respectively.

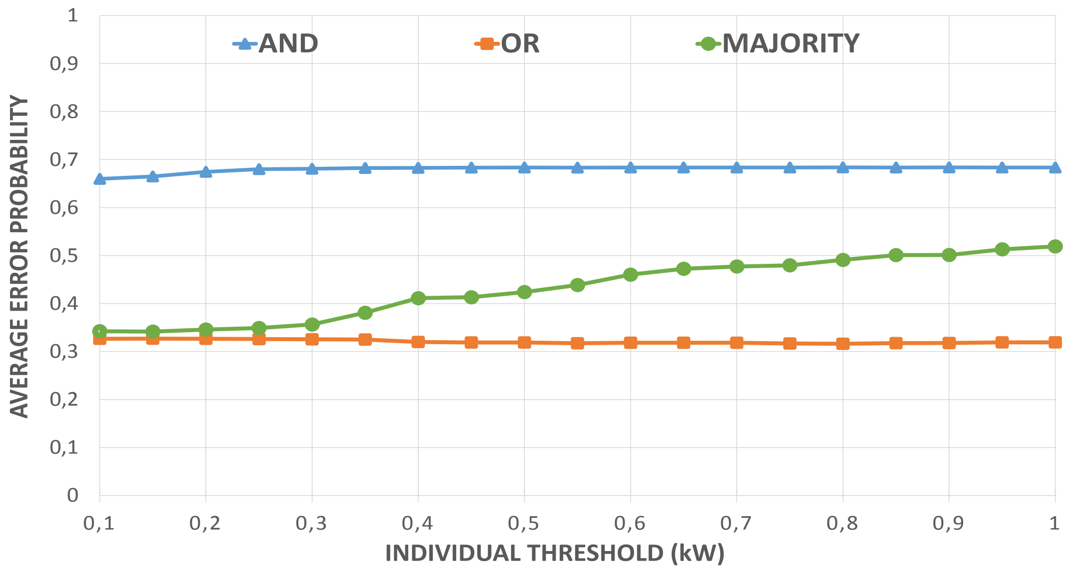

Figure 26: Average Error Probability P_{er} as a function of the individual threshold γ assuming AND, OR and MAJORITY decision rules for $p_1 = 0.2$, $p_2 = 0.1$ and $P_{\text{th}} = 5$ kW. Each point is obtained using Monte Carlo simulation (10^3 snapshots).

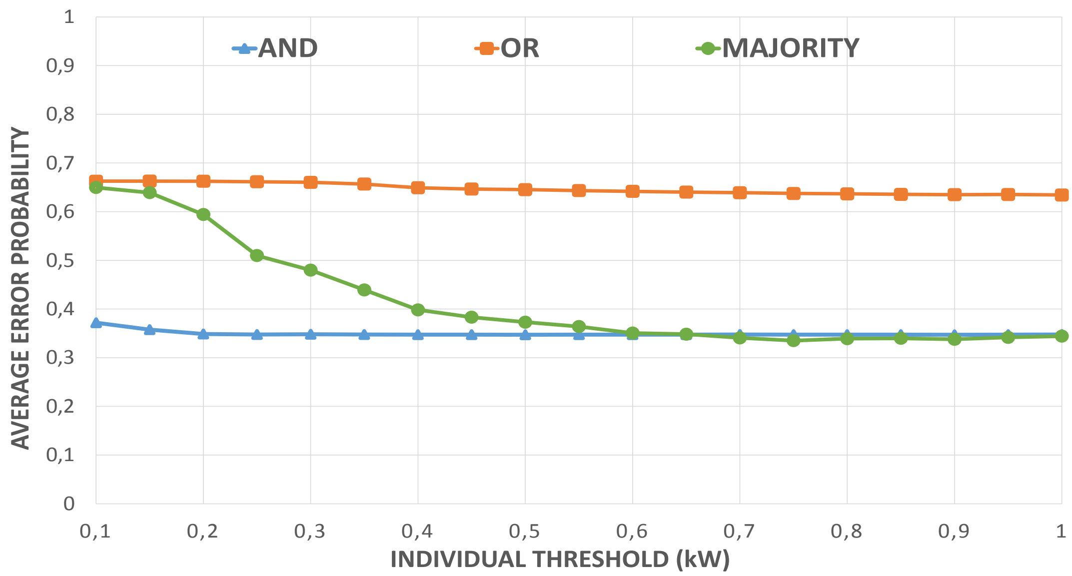

Figure 27: Average Error Probability P_{er} as a function of the individual threshold γ assuming AND, OR and MAJORITY decision rules for $p_1 = 0.2$, $p_2 = 0.1$ and $P_{\text{th}} = 7.5$ kW. Each point is obtained using Monte Carlo simulation (10^3 snapshots).

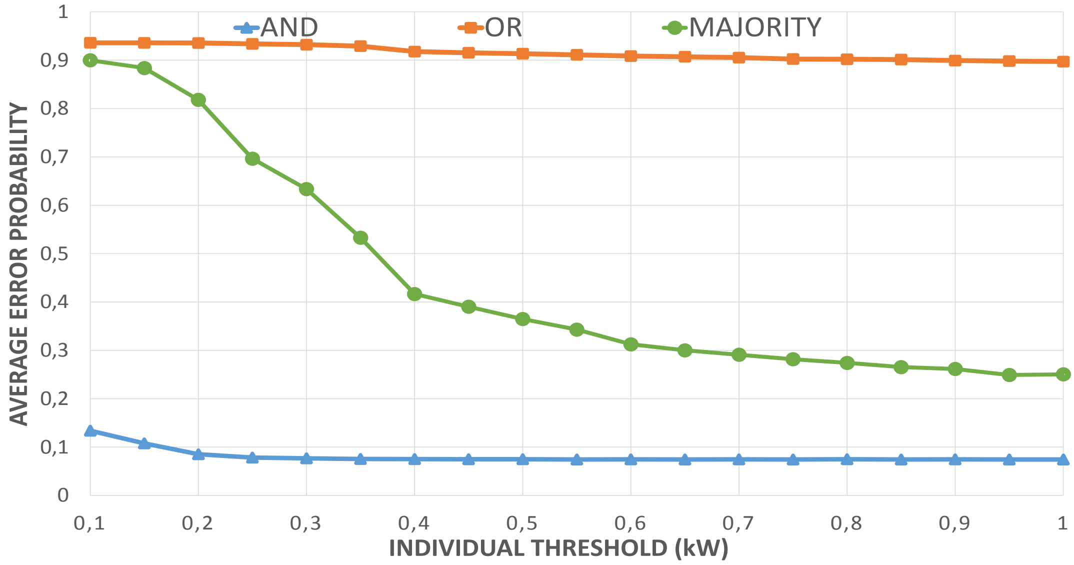

Figure 28: Average Error Probability P_{er} as a function of the individual threshold γ assuming AND, OR and MAJORITY decision rules for $p_1 = 0.2$, $p_2 = 0.1$ and $P_{\text{th}} = 12.5$ kW. Each point is obtained using Monte Carlo simulation (10^3 snapshots).

It can be seen from the results that changing the threshold has a big impact on the average error of the system. When it is set too low (Fig.26) the model with the OR gate works the best. This is due to the fact that OR gate favors 1 due to its nature; Hence, Hence, when the threshold is set with a relatively low value, the signal $s[n] = 1$ is more frequent. and so more ones go into the aggregators input and after that into the network operator. However, the individual threshold γ has little effect on the system performance. When $P_{\text{th}} = 5$ kW, the lowest average error probability is about 30%. This is the case that the OR rule works its best.

Moreover, when the threshold increases to higher values, the results totally change. Increasing the threshold P_{th} modifies this behavior as shown in Figs. 27 and 28 where AND gate starts working better with the lower error probability. Once again, this happens because when the threshold is set higher, the signal $s[n] = 1$ becomes more frequent, meaning that the input signal will below the defined threshold for most of the time slots and thus experiencing more zeros than ones which favors the performance of AND and results in error probabilities $P_{\text{er}} < 10\%$. The parameter γ, once again, has little effect on the error probability for AND and OR. AND rule error probability decreases as the threshold increases higher while OR rule error probability increases with the threshold increasing to higher values.

As for the MAJORITY decision rule, it works most of the times between the other two decision rules since it does not favor *a priori* any state $s[n]$. Since this decision rule will choose the state more frequent, the individual threshold γ will strongly affect its performance. In other words, while AND and OR gates respectively induce $\theta[n] = 0$ and $\theta[n] = 1$, MAJORITY does not induce any state $\theta[n]$. Therefore, although it can have a worse performance, it can be seen as fairer and better represents the system variations. This rule is therefore more susceptible to communication errors

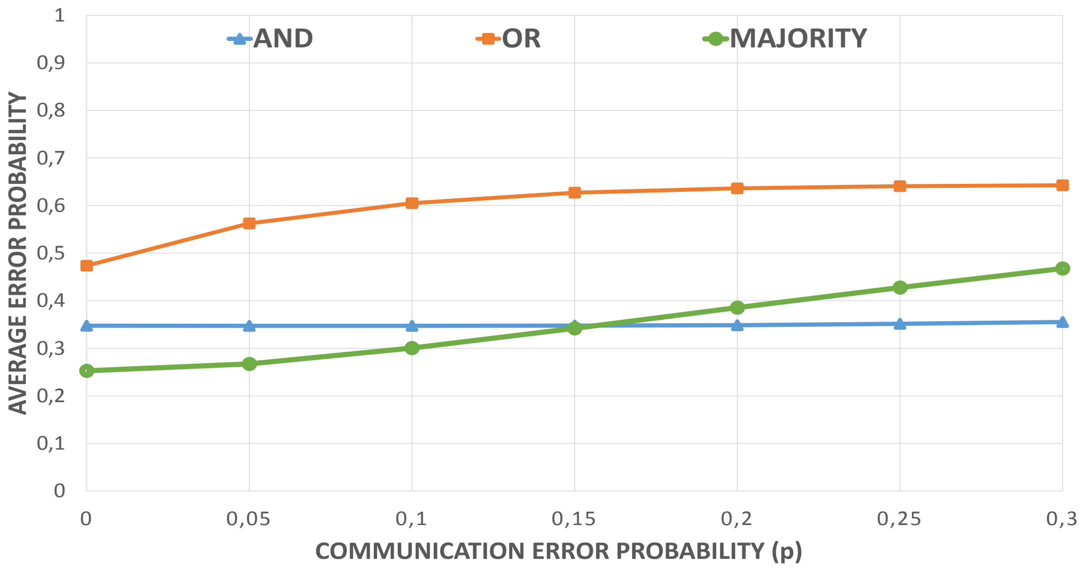

Figure 29: Average Error Probability P_{er} as a function of the communication error probability $p_1 = p_2 = p$ assuming AND, OR and MAJORITY decision rules for $\gamma = 0.6$ kW and $P_{\text{th}} = 7.5$ kW. Each point is obtained using Monte Carlo simulation (10^3 snapshots).

and variations in the individual thresholds γ.

In addition to these results, another interesting observation was the effect of increasing the NAN (p_1) and WAN (p_2) communication error on the average error probability while both of the thresholds are fixed. The results can be seen in Fig.29 . In this figure, it is assumed that $p_1 = p_2 = p$, $\gamma = 0.6$ kW and $P_{\text{th}} = 7.5$ kW.

The first interesting observation from the presented plots is that, even when $p = 0$ (error-free), P_{er} assumes a somewhat high value (about 25%) even in its best case, which is given by MAJORITY. When the communication error p increases, P_{er} also increases for OR and MAJORITY while is kept (approximately) constant for AND. As discussed before, this happens due to the nature of the AND rule, whose performance is determined by the frequency that $s[n] = 0$ occurs and the susceptibility of MAJORITY to more frequent communication errors. It can be said that the MAJORITY rule works most of the times between the other two decision rules but in the case where $WAN_{thre} = 7500$, MAJORITY has the best performance when NAN_{thre} is higher than 600 KW. AND and OR decision rules can be seen as two extreme cases of the MAJORITY decision rule.

Overall, in this chapter, it was observed that it is possible to get a reasonable average error probability if the event $\sum P_{i,j} > P_{\text{th}}$ occurs with low frequency and, therefore, $s[n] = 1$ is rare. In this scenario, AND gate can achieve a low P_{er} since if favors the state $\theta[n] = 0$; the drawback is that by choosing such a rule, the decision is weakly related to the system state. In other words, using AND leads to a quasi-constant guess of $\theta[n] = 0$ (regardless of the error events and the actual individual state) so, as the actual system state is $s[n] = 0$ anyway, the average error probability tends to be low.

MAJORITY rule in turn better captures the system dynamics, but at the same time is much more vulnerable to communication errors.

5. DISCUSSION

In chapter 3 we analyzed a two hop system model where the sensors in houses send their sensing data to their corresponding aggregators. These sensors are conditioned with a specific threshold which depends on different applications of the network. This threshold can be for example power consumption, temperature, etc. If the sensors receive data higher than this threshold they send their data to the aggregators through an error prone binary symmetric channel. The aggregators then send the data they have received to the network operator through another binary symmetric channel which is also subjected to error. Finally the data is received by the network operator which then will decide on the general state of the network.

We observed different error probabilities related to different decision functions AND and OR. The error probability of the AND rule were lowest when the input signal was close to zero at most of the time instants t_n. This indicates that in when the input signal is usually lower than the predefined threshold, using AND rule will result in lower error probability. Having a low error probability is an advantage, specially for applications that we only want to have a general state of the system, for instance, the system operator only needs to know the total power consumption of the system and does not need to know the details of the system or the consumption of the individual elements of the network. In this scenarios, it is very beneficial to use the AND rule when the input signal is usually below the threshold, but if the system operator needs to know the details of the system and each element then because this decision rule is not very sensitive to changes in the input, it is not efficient to use this method.

Then we studied the average error probability of the OR decision rule. We observed that the behaviour of the OR rule is exactly the opposite of the AND rule. OR rule has the lowest error probability when the probability of $S[n] = 1$ happening is higher than the probability of $S[n] = 0$ happening, meaning that when the input signal is above the threshold more frequently, OR rule is the best decision rule to be used by the network operator in case it again needs to know only the general state of the system and does not need detailed information about the individual network elements. Because just like AND rule is not very sensitive to changes in the input signal if it is more below the threshold, OR rule is not very sensitive to the changes in the input signal if this input signal is frequently above the predefined threshold of the system; hence, if we need to know for example the power consumption of the an individual house in one the HANs in the system, it is not suitable to use this model but if knowing the total power consumption of the several HANs in out network is the goal of the analysis, then it is suitable to se this model since it is easy to be implemented and also cheap.

Moreover, examining the effect of increasing the number of sensors (users, smart meters,etc) suggest that the lower bound of the average error probability of each of the studied decision rules gets better. Meaning that the case that has the lowest error probability for each decision rule ($P0-1$ in case of AND and $P1-0$ in case of the OR rule) works even better when the number of sensors increases (The average error probability decreases). However, for the upper bound of the error probabilities, the average error probability increases with increasing the number of sensors and system functionality gets worse. This means that the probability of receiving an error at the output of the network operator if the input signal is above the threshold most of the time in case of AND rule and the probability of receiving an error when the input signal is frequently below the threshold in case of OR rule increases with increasing the number of sensors (Smart meters) in the network.

Also, studying the effect of increasing the sensing error suggests that while the lower bound of the average error probability of each decision rule gets worst with increasing the sensing and communication error (The average error probability gets higher), the upper bound has a different reaction towards this change. When the sensing and communication errors are both more than 50%, the systems actually starts working better in scenarios that it is suppose to have a higher error probability; hence, $P1-0^{AND}$ and $P0-1^{OR}$ starts decreasing when $p \geq 50\%$ and $q \geq 50\%$. These were the results obtained by analyzing the one hop model, we then expanded this theory to the two hop model and observed that in that case too each decision rule is working as expected.

In ch.4 we implemented this theoretical two hop model in a real power consumption model. The model studied in this chapter is slightly different than the one studied in ch.3 since in model the decision making is done at two different stages in the network. In this model both the aggregators in the first hop and the network operator in the second hop are responsible for the decision making. So the data is first processed by the aggregators and then is sent to the network operator. In this model too both the NANa and WAN have communication errors (the effect of sensing error was not considered in this chapter). In addition to AND and OR, MAJORITY decision rule was also studied in ch.4.

Our analysis was based on the average power demand of 12 houses in a 24 hour period which was collected from " The Reference Energy Disaggregation Data Set" (REDD). It was shown that the results obtained in this part agrees with the results of the theoretical analysis of chap.3. We experimented the effect of increasing the threshold of the network on the behaviour of each of the decision rules. The result was that when the threshold of the system is set too low, OR rule has the lowest error probability but increasing the threshold will reverse the reaction of the decision rules and as the threshold gets higher AND rule becomes the most reliable one in terms of the average error probability. MAJORITY rule is always working between the lower and upper bond of the error probability which is set by AND and OR depending on the chosen threshold, although at some point when the threshold is set high enough but not very high and not very low, MAJORITY has the best performance (lowest error probability).

Also, the effect of increasing the communication error in both of NAN and WAN suggested that while OR has the highest error probability, MAJORITY RULE works the best for the first half of the error range and AND rule has the lowest error probability for the second half of the error range. It should be mentioned that in this scenario, the thresholds of the network are constant and $p_{WAN} = p_{NAN} = p$

6. SUMMARY

In this thesis, we analyzed different ways that a wireless sensor network can be implemented using different logical decision rules AND, OR and MAJORITY for a non-critical smart grid application. Our goal was to show whether it is possible to build a low cost communication network using only a 1-bit data signaling. Based on our findings, when the threshold is set too low, indicates that the occurrence of the event 1 is more probable and thus OR rule which favors 1 leads to lower error probability while increasing the level of threshold will cause the occurrence of the event '1' to be rare which will lead to AND rule having lower error probability since it favors event zero. In some cases MAJORITY works even better than AND rule and has the lowest error probability (Fig.27).

We show it is actually possible to attain a low error probability using the proposed scheme if the design parameters are properly set: AND decision rule when the event under consideration is rare. We also pointed out the weakness of this scheme, which favors the more frequent state and it can be seen as always guess that the system is in the more frequent state" decision rule. The MAJORITY rule, on the other hand, better captures the system dynamics while it has the drawback of being more susceptible to error events in the communication.

All in all, although the results presented here have some limitations, it clearly opens up new research possibilities. For example, using different ways of signaling considering more realistic modulations (e.g. Quadrature Amplitude Modulation (QAM)) and channel models is a good way to have a more robust communication system that is simple and easy to implement. Other possibility is to use different decision rules like K-OUT-OF-N, which is a more flexible version of MAJORITY.

Another promising way to develop the proposed framework is to statistically study the average power demand signal. Our idea is to build a signal processing technique that makes use of the signal statistics, which has been shown it is not Gaussian but rather Weibull or Log-Normal [48, 49].

7. REFERENCES

[1] Gao J., Xiao Y., Liu J., Liang W. & Chen C.P. (2012) A survey of communication/networking in smart grids. Future Generation Computer Systems 28, pp. 391–404.

[2] Güngör V.C., Sahin D., Kocak T., Ergüt S., Buccella C., Cecati C. & Hancke G.P. (2011) Smart grid technologies: communication technologies and standards. IEEE transactions on Industrial informatics 7, pp. 529–539.

[3] Farhangi H. (2010) The path of the smart grid. Power and Energy Magazine, IEEE 8, pp. 18–28.

[4] Nardelli P.H., Rubido N., Wang C., Baptista M.S., Pomalaza-Raez C., Cardieri P. & Latva-aho M. (2014) Models for the modern power grid. The European Physical Journal Special Topics 223, pp. 2423–2437.

[5] Wang W., Xu Y. & Khanna M. (2011) A survey on the communication architectures in smart grid. Computer Networks 55, pp. 3604–3629.

[6] Zaballos A., Vallejo A. & Selga J.M. (2011) Heterogeneous communication architecture for the smart grid. Network, IEEE 25, pp. 30–37.

[7] Fan Z., Kulkarni P., Gormus S., Efthymiou C., Kalogridis G., Sooriyabandara M., Zhu Z., Lambotharan S. & Chin W.H. (2013) Smart grid communications: overview of research challenges, solutions, and standardization activities. Communications Surveys & Tutorials, IEEE 15, pp. 21–38.

[8] Sauter T. & Lobashov M. (2011) End-to-end communication architecture for smart grids. IEEE Transactions on Industrial Electronics 58, pp. 1218–1228.

[9] Cisco I. (2009), Internet protocol architecture for the smart grid.

[10] Usman A. & Shami S.H. (2013) Evolution of communication technologies for smart grid applications. Renewable and Sustainable Energy Reviews 19, pp. 191–199.

[11] Lewis F.L. et al. (2004) Wireless sensor networks. Smart environments: technologies, protocols, and applications , pp. 11–46.

[12] Perera C., Liu C.H., Jayawardena S. & Chen M. (2014) A survey on internet of things from industrial market perspective. Access, IEEE 2, pp. 1660–1679.

[13] Ullo S., Vaccaro A. & Velotto G. (2010) The role of pervasive and cooperative sensor networks in smart grids communication. In: MELECON 2010-2010 15th IEEE Mediterranean Electrotechnical Conference, IEEE, pp. 443–447.

[14] Erol-Kantarci M. & Mouftah H.T. (2010) Using wireless sensor networks for energy-aware homes in smart grids. In: IEEE Symposium on Computers and Communications (ISCC), 2010, IEEE, pp. 456–458.

[15] Locke G. & Gallagher P.D. (2010) Nist framework and roadmap for smart grid interoperability standards, release 1.0. National Institute of Standards and Technology , p. 33.

[16] Pullins S. (2009) West virginia smart grid implementation plan. Tech. rep., Technical Report DOE/NETL-2009/1386, West Virginia Division of Energy, National Energy Technology Laboratory, US Department of Energy.

[17] Galli S., Scaglione A. & Wang Z. (2011) For the grid and through the grid: The role of power line communications in the smart grid. Proceedings of the IEEE 99, pp. 998–1027.

[18] Ancillotti E., Bruno R. & Conti M. (2013) The role of communication systems in smart grids: Architectures, technical solutions and research challenges. Computer Communications 36, pp. 1665–1697.

[19] Parikh P.P., Kanabar M.G. & Sidhu T.S. (2010) Opportunities and challenges of wireless communication technologies for smart grid applications. In: Power and Energy Society General Meeting, 2010 IEEE, IEEE, pp. 1–7.

[20] Kuzlu M., Pipattanasomporn M. & Rahman S. (2014) Communication network requirements for major smart grid applications in han, nan and wan. Computer Networks 67, pp. 74–88.

[21] Bera S., Misra S. & Rodrigues J.J. (2015) Cloud computing applications for smart grid: A survey. IEEE Transactions on Parallel and Distributed Systems 26, pp. 1477–1494.

[22] Gungor V.C., Sahin D., Kocak T., Ergut S., Buccella C., Cecati C. & Hancke G.P. (2013) A survey on smart grid potential applications and communication requirements. IEEE Transactions on Industrial Informatics 9, pp. 28–42.

[23] Anderson M. (2010) Wimax for smart grids. IEEE Spectrum 47, pp. 14–14.

[24] Ayanoglu E. (1990) On optimal quantization of noisy sources. IEEE Transactions on Information Theory 36, pp. 1450–1452.

[25] Gubner J. (1993) Distributed estimation and quantization. In: IEEE Transactions on Information Theory, Citeseer.

[26] Sinha A., Chen H., Danu D., Kirubarajan T. & Farooq M. (2008) Estimation and decision fusion: A survey. Neurocomputing 71, pp. 2650–2656.

[27] Castanon D., Teneketzis D. et al. (1985) Distributed estimation algorithms for nonlinear systems. Automatic Control, IEEE Transactions on 30, pp. 418–425.

[28] Ribeiro A. & Giannakis G.B. (2006) Bandwidth-constrained distributed estimation for wireless sensor networks-part i: Gaussian case. IEEE Transactions on Signal Processing 54, pp. 1131–1143.

[29] Ribeiro A. & Giannakis G.B. (2006) Bandwidth-constrained distributed estimation for wireless sensor networks-part ii: unknown probability density function. IEEE Transactions on Signal Processing 54, pp. 2784–2796.

[30] Dogandžić A. & Zhang B. (2006) Distributed estimation and detection for sensor networks using hidden markov random field models. IEEE Transactions on Signal Processing 54, pp. 3200–3215.

[31] Li J. & AlRegib G. (2007) Rate-constrained distributed estimation in wireless sensor networks. IEEE Transactions on Signal Processing 55, pp. 1634–1643.

[32] Leung H., Seneviratne C. & Xu M. (2015) A novel statistical model for distributed estimation in wireless sensor network .

[33] Nielsen J.J. et al. (2015) What can wireless cellular technologies do about the upcoming smart metering traffic? IEEE Commun. Mag. 53, pp. 41–47.

[34] Lu R., Liang X., Li X., Lin X. & Shen X.S. (2012) Eppa: An efficient and privacy-preserving aggregation scheme for secure smart grid communications. IEEE Transactions on Parallel and Distributed Systems 23, pp. 1621–1631.

[35] Proakis J. (2001), Digital communications, new york: Mcgrawhill.

[36] Berrio L. & Zuluaga C. (2012) Concepts, standards and communication technologies in smart grid. In: 2012 IEEE 4th Colombian Workshop on Circuits and Systems (CWCAS), IEEE, pp. 1–6.

[37] Huq M.Z. & Islam S. (2010) Home area network technology assessment for demand response in smart grid environment. In: Universities Power Engineering Conference (AUPEC), 2010 20th Australasian, IEEE, pp. 1–6.

[38] Beigi-Mohammadi N., Misic J., Khazaei H. & Misic V. (2014) An intrusion detection system for smart grid neighborhood area network. In: 2014 IEEE International Conference on Communications (ICC), IEEE, pp. 4125–4130.

[39] Network S.S. Wide area networks (wan) .

[40] Collier S.E. (2010) Ten steps to a smarter grid. IEEE Industry Applications Magazine 16, pp. 62–68.

[41] Amin S.M. & Wollenberg B.F. (2005) Toward a smart grid: power delivery for the 21st century. IEEE Power and Energy Magazine 3, pp. 34–41.

[42] Kassakian J.G., Schmalensee R., Desgroseilliers G., Heidel T.D., Afridi K., Farid A., Grochow J., Hogan W., Jacoby H., Kirtley J. et al. (2011) The future of the electric grid. Massachusetts Institute of Technology, Tech. Rep .

[43] Fang X., Misra S., Xue G. & Yang D. (2012) Smart grid,the new and improved power grid: A survey. IEEE Communications Surveys & Tutorials 14, pp. 944–980.

[44] Nardelli P.H.J. et al. (2015) Average error probability in wireless sensor networks with imperfect sensing and communication for different decision rules. arXiv preprint arXiv:1508.02253 .

[45] Akyildiz I.F. & Vuran M.C. (2010) Wireless sensor networks, vol. 4. John Wiley & Sons.

[46] Atzori L., Iera A. & Morabito G. (2010) The internet of things: A survey. Computer networks 54, pp. 2787–2805.

[47] Kolter J.Z. & Johnson M.J. (2011) Redd: A public data set for energy disaggregation research. In: Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA, vol. 25, Citeseer, vol. 25, pp. 59–62.

[48] Munkhammar J. et al. (2014) Characterizing probability density distributions for household electricity load profiles from high-resolution electricity use data. Appl. Energ. 135, pp. 382–390.

[49] Sajjad I. et al. (2015) Probabilistic generation of time-coupled aggregate residential demand patterns. IET Gener. Transm. Dis. 9, pp. 789–797.