

UNIVERSITY OF CAPE TOWN

**Ubiquitous Mesh Networking:  
application to mobile communication  
and information dissemination in a rural  
context**

by

Richard Maliwatu

A dissertation submitted in fulfilment of the requirements for the  
degree of Master of Science

in the  
Faculty of Science  
Department of Computer Science



November 2014

*“Learn from other people’s mistakes because you may not live long enough to make enough mistakes of your own”*

Nelson Mandela

# *Abstract*

ICT has furthered the social and economic development of societies but, rural African communities have lagged behind due to issues such as sparse population, low household income, a lack of electricity and other basic infrastructure that make it unattractive for telecommunication service providers to extend service provision. Where the service is available, ubiquitous service coverage has not translated into ubiquitous access for individuals because of the associated costs. A community-wide WMN offering VoIP using fixed telephone handsets has been deployed as a viable alternative to the cellular service provider. The effectiveness of this WMN VoIP service springs from the mobile phone usage statistics which showed that the majority of calls made are intra-community.

This dissertation has been an effort towards improved communication and access to information for the under-served communities. Key contributions include, mobile VoIP support, translation gateway deployment to make textual information accessible in voice form via the phone, IP-based radio for community information dissemination. The lack of electricity has been mitigated by the use of low-power devices. In order to circumvent the computational challenges posed by the processing and storage limitations of these devices, a decentralised system architecture whereby the processing and storage load are distributed across the mesh nodes has been proposed. High-performance equipment can be stationed at the closest possible place with electricity in the area and connectivity extended to the non-electrified areas using low-power mesh networking devices. Implementation techniques were investigated and performance parameters measured. The quality of service experienced by the user was assessed using objective methods and QoS correlation models. A MOS value of 4.29, i.e. very good, was achieved for the mobile VoIP call quality, with the underlying hardware supporting up to 15 point-to-point simultaneous calls using SIP and the G.711 based codec. Using the PEAQ algorithm to evaluate the IP-based radio, a PEAQ value of 4.15, i.e. good, was achieved. Streaming audio across the network reduces the available bandwidth by 8Kbps per client due to the unicast nature of streaming. Therefore, a multicast approach has been proposed for efficient bandwidth utilization. The quality of the text-to-voice service rendered by the translation gateway had a PESQ score of 1.6 i.e. poor. The poor performance can be attributed to the TTS engine implementation and also to the lack of robustness in the time-alignment module of the PESQ algorithm.

The dissertation also proposes the use of the WMN infrastructure as a back-haul to isles of WSNs deployed in areas of interest to provide access to information about environmental variables useful in decision making.

# Biography

The author was born in Zambia, in a remote rural area called Mporokoso where his parents served as teachers at the time. He later continued growing up in other parts of the country and beyond. He read for his CompTIA Network+, CCNA, MCSE at North Carolina State University/Computer Training Unit in Raleigh, NC. He did his BSc Computer Science at the Copperbelt University in Kitwe, Zambia, and BSc (honours) specialising in Information Technology at the University of Cape Town. In between "*computer sciencing*", he practices Karate and plays guitar but, that is not to say he is good at it.

# *Acknowledgements*

Glory be to God!

First and foremost, I would like to thank Dr. Antoine Bagula for proposing the project and giving me the freedom to run off with the ball.

Secondly, I thank my family: Martin, Jane, Barbara, Charity, Harriet, Eddie and Mama for their selflessness, always wanting the best for me and giving me all the support giveable.

I would also like to thank Dominic Follet-Smith for telling me, "*lets cross the Jordan of masters*" and my other friends within the ISAT/ICT4D research group, Taariq Mullins, Jarred Martins, Lutando Tobi Ngqakaza for giving me the inspiration and help whenever I needed it.

Furthermore, my life at UCT would've been miserable if it hadn't been for my karate buddies: Thuso Simon, Katie Hamilton, Rene Schieritz, Bokang Mauba, Alphonie Raphadu, Beatrix Williams and everybody else.

Lastly but not least, I would like to thank the UCT Postgraduate Funding Office and the NRF for awarding me the scholarship which helped me in a big way. I look forward to giving back to my community with a continued attitude of gratitude.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Biography</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>List of Figures</b>	<b>viii</b>
<b>List of Tables</b>	<b>x</b>
<b>Abbreviations</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Aims of the dissertation . . . . .	3
1.2 Research questions . . . . .	3
1.3 Motivation . . . . .	4
1.4 Contribution and organisation of the dissertation . . . . .	5
<b>2 Background</b>	<b>7</b>
2.1 Information needs of rural communities . . . . .	7
2.1.1 Information needs of small scale farmers . . . . .	8
2.2 Enhancing access to information . . . . .	10
2.2.1 Interactive Voice Response . . . . .	10
2.2.2 Telephone accessible website . . . . .	11
2.3 Wireless Mesh Networks . . . . .	12
2.3.1 WMN application scenarios . . . . .	13
2.3.2 Integration of WMN with other networks . . . . .	13
2.4 Related work . . . . .	15
2.4.1 Sensor technology in agriculture and environmental monitoring . . . . .	16
2.4.2 Mobile phone communication via wireless mesh network . . . . .	18
2.4.3 Voice based information systems . . . . .	20
2.5 Challenges when supporting mobile devices using WMN technology . . . . .	21
2.6 Conclusion . . . . .	22
<b>3 System design and implementation</b>	<b>23</b>
3.1 Design motivations . . . . .	24

3.2	System framework . . . . .	25
3.2.1	Users . . . . .	25
3.2.2	Monitoring . . . . .	26
3.2.2.1	System requirements . . . . .	27
3.2.3	Information dissemination and communication . . . . .	28
3.2.3.1	Voice-to-voice, text-to-voice community information system . . . . .	29
	Regular user usage scenario: . . . . .	30
	Operator usage scenario: . . . . .	31
3.2.3.2	Organisation of information . . . . .	31
3.2.3.3	Interactive Voice Response (IVR) design considerations . . . . .	32
3.2.3.4	Translation gateway . . . . .	34
	Translation process . . . . .	35
3.2.3.5	IP-based community radio . . . . .	36
3.2.3.6	Mobile communication using wireless mesh network . . . . .	37
3.3	Modifications needed on the WMN infrastructure to support new services . . . . .	39
3.3.1	Supporting mobile users . . . . .	39
3.3.2	Optimizing performance . . . . .	40
3.4	Conclusion . . . . .	41
<b>4</b>	<b>Performance evaluation</b> . . . . .	<b>42</b>
4.1	WMN test-bed design . . . . .	43
4.1.1	Node description . . . . .	43
4.1.2	Enabling multi-hopping on WMN test-bed . . . . .	45
4.2	Determining the mobile VoIP call quality . . . . .	45
4.3	Determining VoIP capacity of mesh node . . . . .	49
4.3.1	Experimental methodology . . . . .	50
4.3.2	Results of VoIP capacity of mesh node . . . . .	50
4.4	Testing the IP-based audio streaming . . . . .	52
4.4.1	Findings . . . . .	53
4.4.2	Evaluation of IP-based radio using PEAQ algorithm . . . . .	53
4.5	Testing the text to voice translation gateway . . . . .	56
4.5.1	Evaluation of the synthesised speech using PESQ . . . . .	57
4.5.2	Some limitations of the translation gateway . . . . .	59
4.6	Summary and implications of the results . . . . .	60
<b>5</b>	<b>Discussion on the findings</b> . . . . .	<b>61</b>
5.1	How can the WMN be used for mobile VoIP? . . . . .	61
5.1.1	Factors affecting the quality and capacity of mobile VoIP . . . . .	63
5.2	How can the WMN infrastructure be used to disseminate community information? . . . . .	65
5.2.1	Can the services be implemented solely on low-power devices? . . . . .	67
5.3	What is the impact on overall network performance of adding the new services on the WMN? . . . . .	68
5.3.1	Recommendations . . . . .	69
<b>6</b>	<b>Conclusion</b> . . . . .	<b>71</b>
6.1	Future work . . . . .	73

<b>A Appendix</b>	<b>74</b>
A.1 Sample measurements . . . . .	74
A.2 Configuration and code snippet . . . . .	77
<b>References</b>	<b>80</b>



# List of Figures

1.2	Mobile-cellular subscriptions by region, total and per 100 inhabitants, 2005-2014. Note: * estimates. Source: ITU	3
2.1	Illustrating the WMN architectures.	14
2.2	Connecting the VoIP network to PSTN.	15
3.1	Components of a drought early warning system. Source: adapted from a framework for predicting droughts in developing countries using sensor networks and mobile phones [1].	23
3.2	Mobile communication and information dissemination using wireless mesh network system architecture. Inspired by DFAS [2].	25
3.3	Using the wireless mesh network as back-haul to islands of wireless sensor networks.	27
3.4	Rural community information system.	29
3.5	IVR design model.	33
3.6	High level view of the translation gateway.	34
3.7	Text to voice translation process.	35
3.8	Community IP radio station audio streaming overview.	36
3.9	Using wireless mesh network for mobile communication. Source: adapted from the SIP trapezoid [3].	38
4.1	Alix system.	44
4.2	VoIP stress test setup.	51
4.3	Audio streaming on multi-hop single-path network configuration.	52
4.4	Audio streaming on multi-hop multi-path network configuration.	52
4.5	Audio streaming bandwidth utilisation based on a 1Mbps link capacity.	53
4.6	Basic principle behind objective measurement of audio quality.	54
4.7	Investigating the effects of network hop count on audio quality using ogg audio format at 44kHz sampling rate.	55
4.8	Investigating the effects of sampling rate on audio quality caused by the common codecs.	56
4.9	Assessing the quality of automatically generated speech.	59
5.1	Distributed-centralised mobile VoIP architecture.	63
5.2	Unicast audio streaming bandwidth utilisation.	70
5.3	Illustrating bandwidth utilisation for audio streaming using multicast.	70
A.1	Mean CPU utilisation given the number of active voice channels.	75
A.2	Memory usage with active voice channels.	75

---

A.3	Basic sip configuration used to setup sip clients. . . . .	77
A.4	Sample <i>call file</i> used to automate calls or text messages. . . . .	77
A.5	Basic extension configuration, determines what to do with a call when received at the PBX. . . . .	78
A.6	Automatic mapping of textual information to speech. The resulting audio files are then made accessible via a phone service. . . . .	79

# List of Tables

2.1	Small scale farmers' issues and potential solution. . . . .	9
4.1	E-model output scales and definition of categories of speech transmission quality. Source: ITU-T Recommendation G.107 [4] . . . . .	46
4.2	Advantage factor ( $A$ ) example values. . . . .	48
4.3	Coefficients for the respective codecs . . . . .	48
4.4	Codec factor values. . . . .	48
4.5	ITU-R five grade impairment scale in reference to audio quality measurement. . . . .	54
4.6	MOS scale categories . . . . .	58
A.1	Audio streaming PEAQ scores for the common codecs. . . . .	76
A.2	MOS-LQO scores using PESQ method to evaluate the text to audio translation gateway. . . . .	76

# Abbreviations

<b>AP</b>	Access Point
<b>DTMF</b>	Dual Tone Multiple Frequency
<b>GSM</b>	Global System for Mobile
<b>ICT</b>	Information Communication Technology
<b>IP</b>	Internet Protocol
<b>ITU-T</b>	International Telecommunication Union -Telecommunication Standardisation Sector
<b>IVR</b>	Interactive Voice Response
<b>MAC</b>	Media Access Control
<b>MOS</b>	Mean Opinion Score
<b>MOS-LQO</b>	Mean Opinion Score - Listening Quality Objective
<b>MP</b>	Mesh Potato
<b>PA</b>	Precision Agriculture
<b>PEAQ</b>	Perceptual Evaluation of Audio Quality
<b>PESQ</b>	Perceptual Evaluation of Speech Quality
<b>PSTN</b>	Public Switched Telephone Network
<b>QoE</b>	Quality of Experience
<b>QoS</b>	Quality of Service
<b>SIP</b>	Session Initiation Protocol
<b>TTS</b>	Text To Speech
<b>VoIP</b>	Voice over Internet Protocol
<b>VSAT</b>	Very Small Aperture Terminal
<b>WiFi</b>	Wireless Fidelity
<b>WMN</b>	Wireless Mesh Network
<b>WSN</b>	Wireless Sensor Network

*Dedicated to the memory of my dad O.K. Maliwatu. The song no longer plays but, the melody lingers on . . .*

# Chapter 1

## Introduction

Setting up telecommunication infrastructure in rural areas comes at a cost that might be hard to justify in terms of the return on investment. This is in part due to sparse population characterising these areas, which results in low user density. The problem is further compounded by a lack of complementary services such as education, electricity, as well as low household incomes resulting in low consumer paying capacity, thereby making it unattractive for mainstream telecommunication service providers to extend service provision to these communities. Delivering ICT services in rural areas is a challenging task due to the absence of networking infrastructure. Among the available alternative technologies, wired networks are unaffordable and sometimes made impractical by the terrain. VSAT is capable of covering remote areas but the associated initial and recurring costs are prohibitive [5]. Therefore, WiFi is considered as a viable solution in delivering communication services in rural areas.

Though cellular network service may be available in some of the rural communities, the cost of communication remains beyond the means of large population segments. This is evidenced by mobile technology uses such as *beeping* or *buzzing* (i.e. intentional missed calls) that people have innovated to circumvent the cellular network costs [6]. The cellular network service providers tend to charge uniform rates across all subscribers regardless of income potential thereby disadvantaging the low income households. Therefore, it may be said that ubiquitous service coverage has not translated into ubiquitous access. This has been a major driving force behind innovations for alternative communication technology, especially alternatives which do not require licenses to transmit. Voice over Internet Protocol (VoIP) is one alternative which can be implemented on a free band based network thereby not requiring spectrum licensing.

Despite the high cost, the penetration of mobile cellular networks makes it easier for people in rural areas to more readily identify with voice services than they would Internet, email or other non-voice messaging systems. In addition, the high levels of textual illiteracy makes voice communication the most effective form of communication.

Community wireless networks proffer rural community members a means of communication independent of the costly and sometimes unavailable mainstream telecommunication operators. The effectiveness of community networks lies in the fact that the bulk of communication in these areas is intra-community. This is in line with studies which found that most of their phone calls are made to someone within the community [7].

Despite the popularity of the Internet in urban areas, it is something the majority of the rural population is not able to identify with. This might be explained by the low Internet penetration in the African region as shown in figure 1.1. However, voice based communication is something they are familiar with. Therefore, the WMN deployed in the community is primarily used for Voice over IP (VoIP) communication. The people are familiar with voice-based communication because when technology is put aside, voice is the most natural way of communication. Secondly, the proliferation of mobile phones as evidenced by figure 1.2 makes it easy for people to readily identify with alternative voice-based communication technologies. In addition, a voice-based information system enables the provision of information where Internet or computers might not be available, which is the case in most rural areas of Africa.

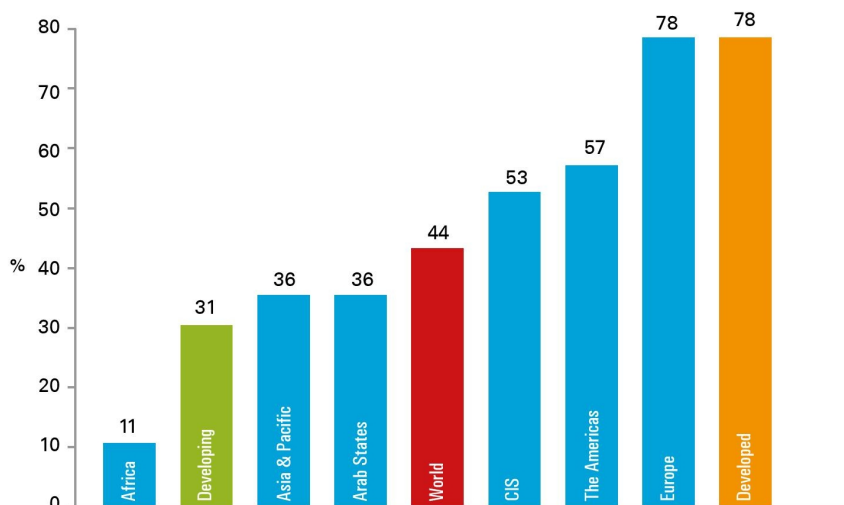


FIGURE 1.1: Percentage of households with Internet access, by region, 2014. Source: ITU <sup>1</sup>.

<sup>1</sup><http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf>

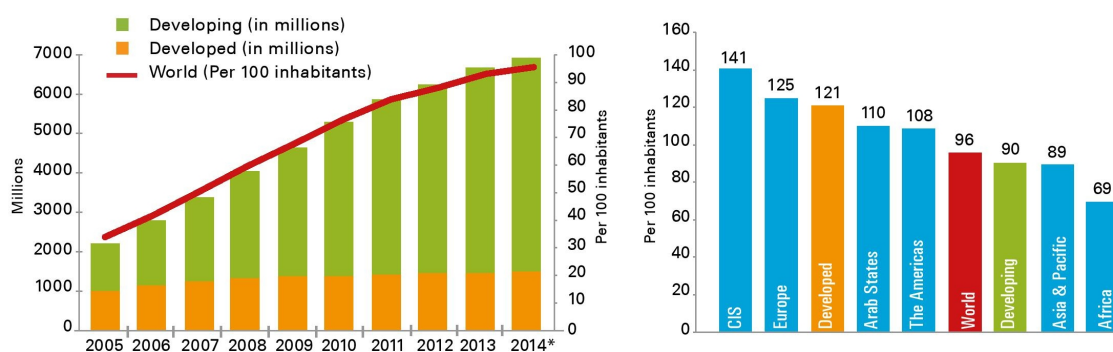


FIGURE 1.2: Mobile-cellular subscriptions by region, total and per 100 inhabitants, 2005-2014. Note: \* estimates. Source: ITU

## 1.1 Aims of the dissertation

While the infrastructure-oriented approach to service delivery has been successful in urban and high-income communities, it is evident that rural communities have been inadequately served. This dissertation builds upon the current state of WMN technology to deliver different services to different users in isolated and under-served areas. The current VoIP network consists of static or wired telephone handsets and so, the aim was to investigate provision of support for seamless mobility so that mobile phones can be used for VoIP calls as well. Mobile phones are a well known technology and once mobile phones connect to the network, the aim was to provide additional voice-based phone accessible services on the WMN. The objective was to ensure that the new services are added without disrupting the primary service. This was achieved by analysing the resource consumption of the new services at both node level and network level so as to ascertain service implementation using the available infrastructure as well as low-power devices.

## 1.2 Research questions

The research focused on adding mobile support to the VoIP service and improving information access and dissemination using a WMN. In addition, information about environmental conditions has the potential to aid small-scale farming communities in their decision making process. Therefore, it is desirable to provide ubiquitous access to this information and present the information in formats appropriate for users with low textual literacy. The study aimed to tackle the following specific questions and sub-questions:



- How can the wireless mesh network be used for mobile VoIP?
  - What are the factors affecting the quality and capacity of mobile VoIP?
- How can the WMN infrastructure be used to disseminate community information?
  - Can the services be implemented solely on low-power devices?
- What is the impact on overall network performance of adding the new services on the WMN?

### 1.3 Motivation

The uniform service charge imposed by cellular network providers puts the rural population at a disadvantage because of their low income potential. This dissertation was motivated by the importance of voice communication over distance and the need to make the service more accessible. Alternative infrastructure and service models are required in order to extend the benefits of ICT to under-served areas. Numerous other studies [8–11] have reported the impacts that information and communication technology has had on rural and development marginalised communities. Some of the potential benefits include:

- Better decisions resulting from information availability and ease of communication.
- Improved healthcare.
- Increase in crop yield due to enhanced access to information on farming practices and market information.
- Political freedom.
- Improved governance.
- Increase in income as a result of increased productivity.
- Introduction of other income streams for predominantly small-scale farming population.
- Economic growth and poverty reduction as a result of a combination of the above listed benefits.

Furthermore, the use of sensor technology in agriculture is known to improve crop yield and quality [12]. However, it has been observed that existing agricultural monitoring

---

systems are mostly applied and utilized in closed agricultural environments such as greenhouses, cattle sheds, etc., as it is difficult to apply agricultural monitoring systems in outdoor locations such as paddies, orchards, etc. because of a lack of IT infrastructure [13]. The proposed system if applied to outdoor agricultural environments is envisaged to contribute towards increased crop yield and improved quality by supplying farmers with information to help them make decisions.

## 1.4 Contribution and organisation of the dissertation

In order to overcome the challenge associated with presenting information to users with limited textual literacy, this dissertation proposes the inclusion of translation gateway(s) on the WMN that translate textual information such as processed data from sensor readings into voice messages to be delivered to small-scale farmers. The use of sensor technology in agriculture is known to improve crop yield and quality. Widespread adoption has only been hampered by a lack of infrastructure. Information gathered from sensors deployed in areas of interest to gather environmental data aids in decision making. This dissertation makes a theoretical contribution on the application of WMN towards environment monitoring and proposes the use of the WMN infrastructure as a back-haul to isles of wireless sensor networks deployed in areas of interest.

Besides the need to monitor the environment and quantify variables for purposes of informed decision making, communication is considered as a basic human necessity. One of the United Nations Millennium Development Goals <sup>2</sup> addresses the need to make available the benefits of new technology especially information and communication technology. The inspiration for this study came in part from the realisation that wireless mesh networks were being deployed and used for VoIP communication in rural communities such as Mankotsi in rural Eastern Cape of South Africa. The network comprises Mesh Potato (MP) from Village Telco <sup>3</sup>. The MP acts as both a network client, sending and receiving data, and as a router, participating in the relaying of traffic across the network. The purpose of this study was to ascertain the implementation of additional services aimed at enhancing communication and access to information using the existing infrastructure. The term *static* has been used to refer to the fixed telephone handsets plugged into the MP which are used for VoIP calls, whereas *mobile* refers to the use of mobile phones connected through an Access Point to access network resources.

The remainder of this dissertation is organised as follows. Chapter 2 provides further motivation for the study and presents the background of the underpinning concepts

---

<sup>2</sup><http://www.un.org/millenniumgoals/pdf/report-2013/mdg-report-2013-english.pdf>

<sup>3</sup><http://villagetelco.org/mesh-potato/>

employed in the design of the system framework. The chapter gives the context of the study and overview of related work. Chapter 3 focuses on the system design, design motivation and provides the implementation details of the system components. Chapter 4 describes the test-bed and the tests carried out to assess the quality of each of the services. Chapter 5 discusses the obtained results and revisits the research questions answered through the output of this dissertation. Chapter 6 presents the conclusions drawn from the study and outlines plans for further work.

## Chapter 2

# Background

The surrounding ideas of this dissertation draw from several fields of study which may broadly be categorised as wireless mesh networking, wireless sensor networks, communication and voice based information systems in a rural context. The wireless mesh network is considered as a viable low cost alternative infrastructure to cellular networks for community wide communication. This is sometimes referred to as *inverse infrastructure* [14]. The term inverse infrastructure describes systems not owned by government or large corporations and are not centrally controlled which is in clear contrast with the way telecommunication has traditionally been supplied in the past. Inverse infrastructures are user-driven, bottom-up development based on self-organisation. On the other hand, a wireless sensor network consists of sensors which are able to sense and transmit data regarding various environmental variables of interest. Sensors are highly application specific. In this dissertation the main concern is with sensor data obtained over a short to medium term period. Furthermore, a voice-based information system is a system which does not require reading or writing capabilities. Navigating for information through the system is via voice prompts and pressing specified keys. The rationale behind the voice-based approach is discussed in greater detail in section 2.2.

### 2.1 Information needs of rural communities

Understanding the information needs of a community requires involvement and interaction with the community for a sufficient period of time in order to gain knowledge of the kind of questions people might have. According to the literature surveyed [15], [16], [17], [18], the information needs of rural communities are non-uniform but may be placed in eight broad categories summarised below:

- **Education:** pertains to school going individuals, either primary or secondary or those who might be doing distance education.
- **Government:** information about Government policies and development plans for the community.
- **Agriculture:** this category encompasses agriculture related issues such as advise to farmers or information on best farming practices, market prices for farm produce, sources and prices of farm inputs such as fertilizer, seeds, etc.
- **Health:** rural areas tend to be characterized by a sparse distribution of health centres. Physicians tend to visit these communities every now and then. Therefore, information such as physician schedule and other health related information need to be made available.
- **Business:** every now and then someone might have something which they wish to buy or sell. Information regarding items being sold or merchandise required by potential buyers has to be made available.
- **People:** this section pertains to information about people such as contact details of people with special skills or trades.
- **Community news:** communities tend to have meetings from time to time organised by community leaders. Members of the community need to know when and where these meetings are. In addition, minutes of the meetings can be kept to serve the community.
- **General:** The general category can be used for information that does not fit perfectly in any of the prescribed categories. This could be static or quasi-static community information such as population, school related, etc.

In addition, because of the mobile phone penetration, an ideal information system is one that is voice and mobile phone based because it eliminates the need for users to learn new modes of interaction or a new technology.

### 2.1.1 Information needs of small scale farmers

The agricultural system in a number of African countries consists of extension officers whose role is to give support to small scale farmers in terms of agricultural expertise. However these extension officers are not able to reach all the farmers because of transport logistics and also because the number of farmers is larger in relation to that of the extension officers. This has led to the need for innovative services to deal with the gap.

The services that have been innovated may broadly be classified as voice-based and web-based information services.

One extension officer was quoted as saying, *"We asked our government for bicycles, and we got them, but we could not reach all our farmers. Then we asked for motorcycles, and we got them, but we still could not reach all our farmers. Now we have radio, and through radio we can reach all our farmers"* ([19], page 1).

The above statement underscores the challenges faced by extension officers in their efforts to try and reach the farmers. Extension officers require more effective alternative ways of supplying farmers with the information they need.

Furthermore, the existing services are often top-down, leaving the local farmer uninvolved in the processes. This results in an imbalance between generic information and locally relevant information. While an ideal information system for the small scale farmer in rural areas may not provide the other needed structures such as financial systems or other institution, it can assist in providing information for them to better use such establishments.

Table 2.1 summaries the information system issues faced by small scale farmers in rural areas and how the proposed system attempts to alleviate the situation.

<b>Issue</b>	<b>Solution</b>
Extension officers are unable to reach all the farmers due to capacity constraints	By creating an information dissemination platform, extension officers can reach more farmers
Lack of interaction	Having a cost effective distance communication system will encourage interaction
Local knowledge is over looked	Allowing the users to upload questions or comments and browse through what other users have posted ensures that the content is kept relevant
No computer or Internet access	Using a phone based information system since phones are readily accessible
Farmers with limited textual literacy are unable to substantially benefit from ploughing through Web pages or printed media	A voice based system is such that even users with limited literacy are able to use

TABLE 2.1: Small scale farmers' issues and potential solution.

While a voice based solution has the advantage of potentially reaching even textually illiterate users in that there is no reading capability required, the system may not provide detailed information as is achievable with a pictorial illustration on a web page for example. However, the ease with which the language can be customised and the natural

relationship users have with voice communication makes a phone based solution the most appropriate for small scale farmers.

## 2.2 Enhancing access to information

The Internet is replete with a wealth of information. However, not everyone is able to access this information for various reasons ranging from physical disability to lack of Internet connectivity.

There has been extensive work in the recent past aimed at enhancing access to Web-based information. For example, the Web Accessibility Initiative of the World Wide Web Consortium's has created guidelines for interface design to primarily address issues affecting the physically disabled individuals [20]. The purpose of the guidelines is to ensure that people with disabilities such as visual or hearing impairment, limited motor control or cognitive ability, etc. have equal access to the Web. This is achieved by firstly, eliminating barriers to access through provisions such as alternative input modalities such as speech in addition to the conventional keyboard and mouse as well as support for assistive technologies which mimic the conventional input devices. Secondly, information access is enhanced by providing the output in a number of ways or formats such as text or audio depending on the user scenario. Just as having information in audio format may aid the visually impaired people, having information in audio format such as podcasts transcribed into text aids those with hearing disabilities.

### 2.2.1 Interactive Voice Response

Voice-based services allow users to access information by initiating a call to a service using a given service number. These voice-based systems are referred to as Spoken Dialogue Systems (SDS) or Interactive Voice Response (IVR) systems [21]. Users interact with an SDS for both input and navigation through speech. On the other hand, IVR users interact with the system using Dual Tone Multiple Frequency (DTMF). The playing of recorded messages or prompts combined with input collection through key presses characterises the IVR. Using IVR can be helpful in automating service provision. The IVR has the advantage of allowing the services to be offered round the clock and the elimination of a human personnel reduces the cost of service provision.

The most challenging human characteristic to consider in IVR design is the cognitive capabilities of the expected users of the system [22]. This is due to the fact that quite often the target user population is a general population, with its inherent range of

---

cognitive abilities. Difficulties arise from differences in the users expected task flow, the language model used to describe the flow, and the vocabulary with which the IVR is presented. Therefore, some of the considerations to make when designing an IVR are the language and persona. The term persona is being used in this context to refer to the nature of the prompts associated with the IVR such as the accent, tone and intonation of the voice prompts. The voice prompts have to be in a language understood by users and the persona has to be carefully designed to suit the users of the system.

In some cultures <sup>1</sup>, the choice of a male or female voice prompts can have a significant impact on the users' impression of the system. For instance, in cultures where men do not customarily take instructions from women, it would be very difficult for male users to take advise information from the system if the prompts are driven by a female voice.

When the demographics dictate the use of a multilingual IVR several decisions need to be made such as the number of languages and which ones to provide for. In addition, answers to the questions, *how do you put the various languages together, at what stage in the IVR do you provide the language options and in what order?* need to be provided.

One approach is to customize the language by profiling the users based on the caller ID [23]. This would work accordingly where there is a strict one phone one user ownership model. However, in developing countries another type of ownership model has been identified where a single phone gets used by several individuals [24]. In this type of situation, the best approach is to provide an upfront language menu in all the languages or use dedicated different service numbers for each of the languages being catered for. Using a multilingual IVR might be essential where the system covers a diverse demography. In the South African context, a multilingual IVR aids in allowing users to exercise their linguistic constitutional right.

### 2.2.2 Telephone accessible website

In tandem with the principle of universal access, telephone accessible websites have been developed to give people access to Web-based information who might not have a Web browser or Internet connectivity but have access to a telephone.

Telephone accessible websites use the VoiceXML technology to enable users navigate the site by telephone. VoiceXML is an application of the eXtensible Markup Language (XML) defined by the World Wide Web Consortium (W3C). It is a domain-specific

---

<sup>1</sup>CNN published an article where they pointed out that BMW was forced to recall a female-voiced navigation system on its 5 Series cars in the late 1990s after being flooded with calls from German men saying they refused to take directions from a woman: <http://edition.cnn.com/2011/10/21/tech/innovation/female-computer-voices/>



---

language that defines dialogs between humans and machines in terms of audio files to be played, text to be spoken, speech to be recorded or recognized, and touch-tone input to be collected [25].

The basic framework comprises a VoiceXML gateway which logically sits between the web server and the telephone network. The VoiceXML gateway maps a designated telephone number to a URL. Users interact with the web server via the gateway. Depending on the implementation, users may interact with the system by speaking if speech recognition is supported or by pressing keys on the touch-tone where DTMF is supported. At the core of the VoiceXML-based implementation are two key speech technologies namely *speech synthesis* and *speech recognition*. This implies that the implementations are subject to the inherent limitations of the underlying technologies such as the lack of naturalness in the synthesised speech and inaccuracies in speech recognition.

Despite all these enhancements, information access for the rural population is yet to be fully addressed. Rural communities of Africa face added challenges such as limited textual literacy, lack of relevant content, lack of electricity and other basic infrastructure.

## 2.3 Wireless Mesh Networks

Because of their many advantages such as low upfront costs, robustness, long distance coverage, ease of setup, expansion and maintenance, self configuration, Wireless Mesh Networks (WMN) are seen as a viable approach towards improving communication in rural and under-served areas [26], [27], [28]. A WMN is comprised of mesh routers and mesh clients. The term client is being used to refer to end devices such as laptops, sensors, etc. Clients conventionally connect to routers and in turn routers connect to other routers thereby forming a multi-hop mesh architecture so much so that a single point of failure is prevented. Mesh routers also tend to connect to other networks and act as gateways. These networks may be heterogeneous and support various services. There are several WMN architectures but according to Akyildiz[29], they may be placed in one of the three general categories depending on the functionality of the nodes:

- *Infrastructure WMN*. Routers form the mesh with possibly multiple radio technologies and perform the task of relaying packets from source to destination. Clients i.e. end devices, connect to the infrastructure through Mesh Access Points as shown in figure 2.1(a).
- *Client WMN*. Clients connect to other clients which in turn connect to other clients thereby forming a mesh as illustrated in figure 2.1(b). In addition to providing

end user applications, clients perform the routing and network configuration functionality.

- *Hybrid WMN*. This architecture combines Infrastructure and Client WMN. Clients connect to other networks such as the Internet or sensor network through the mesh routers. Clients are also able to connect to the network by connecting to other mesh clients. This type of architecture in which routing capability is built into the clients extends the WMN coverage.

The ability to extend coverage is among the advantages of WMN. In the majority of network applications, throughput is a key performance factor. However, coverage and throughput enhancement are two opposing goals. This is even more so in wireless networks which tend to suffer interference from other wireless transmissions. An increase in the coverage area decreases the throughput due to longer transmission distances between the nodes. Several frameworks aimed at solving the data routing problem at the network layer and the transmission power issue at the physical layer have been proposed to enhance scalability while optimising throughput. Such proposals include multi-channel Ring-based WMN architecture [30] and cross-layer optimization framework for multi-hop multicast [31].

### 2.3.1 WMN application scenarios

WMNs have been used in several applications such as community and neighbourhood networking, enterprise networking, metropolitan area networks for purposes of sharing Internet connectivity and access to services such as email, and so on. Owing to their characteristics, WMNs have been applied in numerous scenarios such as large-scale security video surveillance systems, health and medical systems, building automation, disaster recovery and rescue, peer-to-peer resource sharing, etc. [32], [33], [34], [28], [35].

### 2.3.2 Integration of WMN with other networks

WMN are integrated with other networks such as the Internet, cellular, sensor networks, etc., using bridges and gateways. While the terms *bridge* and *gateway* may be used interchangeably, in this dissertation the term gateway is being used to refer to a device that allows information exchange between disparate networks. On the other hand, a bridge is much simpler in functionality and might be used for expansion or connection of geographically separate networks.

Conceptually, a gateway can be used to connect the wireless mesh network to the Internet thereby extending the coverage beyond the community. However, in reality this depends

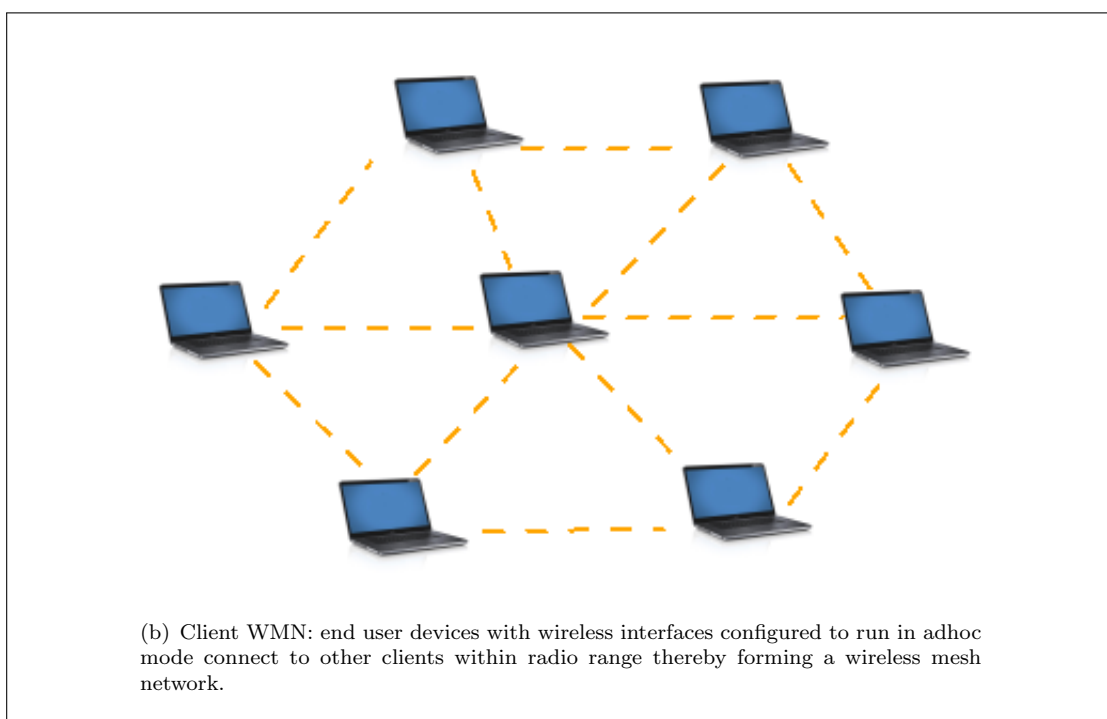
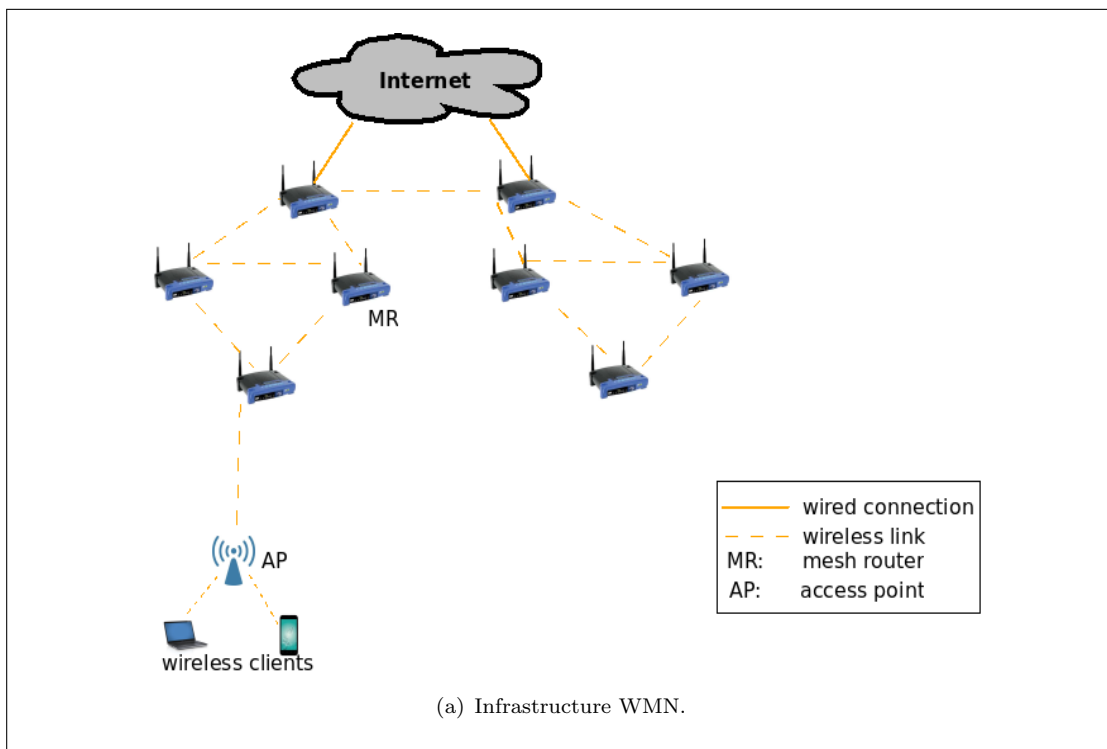


FIGURE 2.1: Illustrating the WMN architectures.

on the sole purpose of the network. For instance the stakeholders of the wireless mesh network deployed in Mankotsi, rural Eastern Cape, indicated that they did not wish to extend connectivity beyond the community [36]. They only wanted to use the network for intra-community VoIP calls. For such a network, one possible desire might be integration with the global telephone system. The integration has to take place at two levels, the physical level and at the logical level.

Asterisk is an open source framework which can be used to build VoIP gateways using an ordinary computer. Asterisk supports connectivity to the Public Switched Telephone Network (PSTN) and thus connecting to the global telephone network. The method of integrating Asterisk with the PSTN depends on whether the telephone carrier uses digital or analogue circuits. For analogue telephony, the gateway has to be fitted with a *Foreign eXchange Office* (FXO) card which allows connection to analogue lines [37]. The FXO is an RJ-11 connector which allows the gateway to connect the private branch exchange (PBX) directly to the PSTN via the Foreign eXchange Subscriber interface (FXS) as illustrated in figure 2.2.

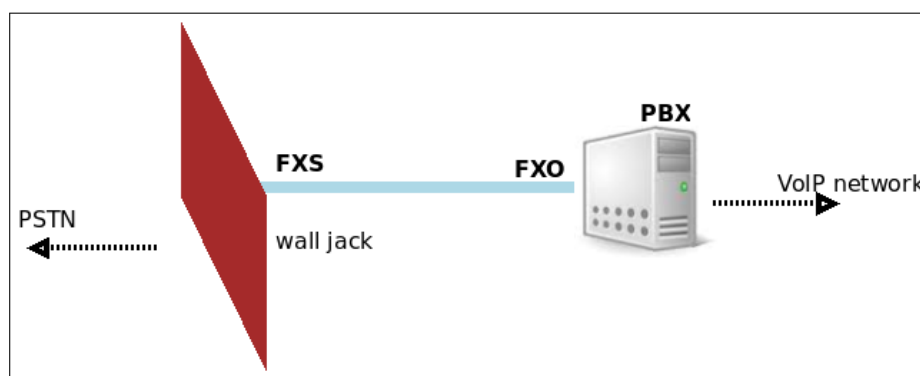


FIGURE 2.2: Connecting the VoIP network to PSTN.

The VoIP server has to then be configured to route PSTN bound calls through the FXO. The FXO card manufacturer usually provides a software framework such as DAHDI<sup>2</sup> needed to enable communication between the PSTN card and the VoIP server.

## 2.4 Related work

There are four main areas of research related to this dissertation namely, VoIP communication, wireless mesh networking, wireless sensor networks in agriculture and voice-based information systems. This section presents existing systems and projects from the academic circle and industry that are closely related to this dissertation.

<sup>2</sup>Digium Asterisk Hardware Device Interface. <http://www.digium.com/en/products/telephony-cards/analog>

### 2.4.1 Sensor technology in agriculture and environmental monitoring

The relevance of using wireless sensor technology in the developing world to enhance research in environment monitoring has been stressed in [38]. The term smart agriculture is generally used to refer to agriculture in which sensor technology is employed. A study conducted by Abdullah and Barnawi [39] reviewed applications of wireless sensor networks in agriculture and identified types of agriculture which might benefit the most from the use of sensor technology. The application of sensor technology in agriculture is aimed at making the tasks of monitoring, data collection and actuation more efficient.

When sensors are used to monitor sites of interest for purposes of customising the management based on the site conditions, it is often referred to as Precision Agriculture (PA). PA comprises a set of technologies which combines sensors, information systems, enhanced machinery, and informed management to optimize production by accounting for variability and uncertainties within agricultural systems [40]. Instead of applying uniform treatment across a wide area, PA is aimed at customising the management practice based on precise site conditions. The first goal might be optimal utilisation of resources for instance, scheduling or site specific irrigation. Mafuta et al proposed in [41] a prototype implementation of an irrigation system that demonstrated the integration of sensor, actuator technologies and sms notification to meet the need for precision agriculture in Malawi. The other goal might be quality control. This type of agriculture is especially applicable in regions and sectors such as the wine industry where demand on high standards of produce justify the use of enhanced machinery to account for variabilities in the farming environment thereby optimising the quality of produce. There is generally a trade-off between yield and quality. Wireless Sensor Network (WSN) technologies have been used to remotely monitor key viticulture parameters in real-time for the purpose of maximising both grape quality and yield [42]. When applied to grape fields, this type of PA is more precisely referred to as Precision Viticulture (PV). Monitoring environmental properties such as soil moisture, pests, etc. enables the farmer to apply the right measures in the right places at the right time.

The other application of WSN in the context of agriculture has been in drought forecasting and drought severity prediction across a large geographical area using distributed WSNs [43]. Weather forecast is conventionally done at meteorological stations by processing large amounts of data about the current state of the atmosphere. The accuracy of the forecast tends to reduce as the forecast range increases in terms of time and location [44]. The use of wireless sensor technology in drought forecast is motivated by among other reasons, the need to collect a large amount of data and make the forecasts more reliable, the low-cost of sensor technology compared to the cost of meteorological

---

instruments and the potential that wireless sensor technology has to do the monitoring using data local to the community.

To overcome the challenges associated with making drought predictions such as its complex natural phenomenon and the high cost of specialized weather forecasting instruments, a framework that incorporates traditional knowledge from communities and data collected from Ubiquitous Sensor Networks into the drought prediction system has been proposed [1, 45]. On the other hand, a Drought Forecast and Alert System (DFAS) has been proposed [2]. DFAS employs wireless sensor networks among other technologies to collect a combination of environmental drought data and MODIS<sup>3</sup> satellite images to achieve timely warning and notification. While DFAS is targeted at relevant departments to aid in their decision making or related resource allocation process, the work does not address the presentation of information to users in rural communities who might benefit from such information if it was made accessible and understandable.

In addition, developing countries are characterised by resource constraints such as sensing equipment and information dissemination systems. In order to deliver the benefits of wireless sensor networks in developing countries, Muthoni and Bagula proposed mobile phone and sensor networks (MobiWSN) drought prediction framework [1], which harnesses the mobile phones' computing power. The hypothesis is based on the smartphones' popularity, processing power and range of wireless connectivity options. Unlike other existing tools that ignore the at risk community and focus on macro level information, the framework places emphasis on the affected community. MobiWSN integrates community elders in the drought prediction scheme who are known to have a wealth of traditional knowledge about predicting droughts by observing their surrounding. These elders also have knowledge on planning and preparing for such droughts. In addition, Muthoni and Bagula proposed attaching sensors to public transport vehicles and non-hostile wildlife such as elephants to aid in collection and transportation of climatic data as they move across the region.

Another field in which the environmental monitoring capabilities the sensor technology has been applied is in pollution monitoring. This includes water pollution monitoring [46] and air pollution monitoring [47]. Pollutant substances resulting from volatile chemical compounds, mining activities and other industrial emissions pose a significant threat to human health as well as the environment. The problem of pollution is yet to be well addressed in the developing world. Gas sensors and other sensors deployed in water sources can be used to monitor water and air pollution in communities thereby providing an early warning mechanism when the pollutant level exceeds the acceptable thresholds.

---

<sup>3</sup>Moderate Resolution Imaging Spectroradiometer. <http://modis.gsfc.nasa.gov/>

---

Differences in the wireless sensor network deployment may be observed between the developed and the developing world. These differences include the choice of communication system for information dissemination, which arises from the differences in the available infrastructure. Owing to these differences Bagula et al. [47] developed Ubiquitous Sensor Network for Development (USN4D), an architecture which factors in, the requirements for deployment in the developing world. These requirements include, openness (i.e. open source software and open source hardware), lower cost of equipment, opportunistic dissemination, multi-layer networking for interoperability and long distance deployment, as well as information localisation.

### 2.4.2 Mobile phone communication via wireless mesh network

In response to the high cost of mobile phone communication, several projects have emerged to try and provide alternatives to the costly mobile service providers. In addition, some rural areas are unserved because they are sparsely populated with low incomes among households thereby making them unattractive for the mobile operators to invest in. Proponents of affordable and available telecommunication for all have proposed the use of alternative infrastructure and approaches to those of the cellular operators. The advantages of wireless mesh networks such as low cost, fault tolerance and suitability for rough terrains or hard to wire environments makes wireless mesh networks a suitable candidate.

Outside of the rural context, the rising popularity of smartphones in metropolitan areas has led to an increase in mobile data traffic. The continued increase in applications such as social media for mobile devices, guarantees an upward trend in the amount of mobile data traffic. Cellular network service providers have to cope with the increase in data traffic on their networks. An overload of the cellular network is likely to reduce the quality of service given to customers. The deployment of WiFi is a common solution which has been proposed to address the problem of cellular offloading [48].

Some of the projects aimed at providing alternative solution to mainstream cellular network service are described below.

**WiBack.** This is an architecture which has been proposed that uses the wireless mesh back-haul networks to provide mobile phone access in rural areas [49]. The architecture integrates nano-Base Transceiver Station (BTS) cells to provide the mobile to air interface. These cells are in turn connected via wireless mesh network. Mobile phones connect to the IP-based wireless network through the GSM air interface mounted on the access points. However, the GSM protocol stack is terminated right at the local network access points thereby allowing open protocols such as Inter-Asterisk eXchange (IAX) or

---

Session Initiation Protocol (SIP) with Real-time Transport Protocol (RTP) to carry the signalling and voice traffic across the network.

The key advantage of WiBACK is that it eliminates dependency and the costs associated with mobile operator infrastructure. GSM nano-cells made of USRP hardware and OpenBTS software allows mobile devices to connect directly to the network without needing any software or hardware modifications. However, the coverage area and the number of simultaneous calls that can be handled is much lower compared to that of conventional GSM cells.

**VillageCell.** Villagecell is another low-cost alternative to cellular networks. The solution uses a Software Defined Radio (SDR) powered by OpenBTS - a software implementation of the GSM protocol stack, and Asterisk - a software Private Branch eXchange (PBX) implementation to integrate GSM with VoIP telephony. OpenBTS provides the GSM to air interface whereas Asterisk performs the call routing. Multiple instances of OpenBTS are deployed to cover the desired area and connection among them is achieved through any IP-based standard technology available such as WiFi [50].

**Serval.** The Serval project aims to supplement the infrastructure oriented approach to mobile communication, which inevitably fails to serve all populations. This project proposes the supplementation of the existing infrastructure oriented approach with a non static infrastructure model to meet the needs of the under-served populations. It has been pointed out that the Mobile Ad-hoc Networks (MANETs), sometimes referred to as client WMN have the potential to address the requirements of the proposed model. These requirements include, no licensing required for spectrum, no carrier needed between caller and callee and no central authority needed for number allocation and other administrative tasks. The proposed approach is to have the mobile device as the only infrastructure that is required [51].

The key advantage of Serval is that it provides potential for infrastructure-less mobile communication. Mobile devices connect with other mobile devices in the coverage area and each of the devices takes part in the relaying of traffic across the network formed on an ad-hoc basis. However, the infrastructure-less approach inevitably requires that the callee be within the radio radius of the caller or sufficient intermediate devices to relay the traffic to the endpoint. This might be difficult to realize in areas characterised by sparse population. Furthermore, a study conducted to evaluate the performance of infrastructure-less VoIP communication concluded that the current peer-to-peer VoIP protocols are not as satisfactory compared to the centralised approach [52].



### 2.4.3 Voice based information systems

Providing information to people in rural areas poses a number of challenges. Firstly, the rural areas of developing countries are characterised by a lack of access to printed media or the Internet. Secondly, there is limited literacy among the majority of the population. In such instances, carrying information as voice enhances the access to information and services. Some of the benefits of a voice-based information system are that minimal infrastructure is required, the service can be accessed via the telephone. Having information accessible via the telephone ensures that the service is available from anywhere and also cuts back on transport related costs, delays and the costs associated with printed media.

Some of the voice based information system projects employing speech technologies which have been undertaken targeting the developing communities are described below.

**Lwazi.** The Lwazi Community Communication Service is an automated telephone-based information service that allows local community managers to disseminate information pertaining to health, employment, social grants, etc. to communities. The service works by allowing a registered municipal communication officer to upload information to the service for the community members through a Web interface. In turn, community members are able to access the information via a telephone service and leave messages for the municipal workers [18].

The Lwazi project makes the South African Government service information available in the 11 official languages of South Africa thereby assisting the Government in service delivery by overcoming the language and literacy barrier.

**Avaaj Otalo.** This is an asynchronous interactive voice application for small scale farmers in India. The system allows users to post questions, listen to others' questions and responses on a number of agricultural subjects such as how to deal with specific crop diseases or pests. The Avaaj Otalo project aims to increase the reach of agricultural extension officers using a voice-based forum to provide access to agricultural knowledge. The service is made available via low-cost mobile phones [53]. However, the service relies on the commercial cellular network whose associated costs might deter some users from using the service.

**VoiKiosk.** This is a content creation and dissemination system by and for users in rural areas. The system is targeted at the rural population where there are no Internet facilities. The system is accessible by phone thereby making it affordable and ideal for the textually illiterate rural population. VoiKiosk consists of a content creation and configuration component which allows operators to upload information in categories

---

ranging from agriculture to health, relevant to the community. Villagers are in turn able to call VoiKiosk for information. By so doing, VoiKiosk acts as an information and service portal for the village. VoiKiosk addresses the problem with the current sources of information such as the majority of newspapers, radio or television programs which is the lack of content relevant to the local community [17].

## 2.5 Challenges when supporting mobile devices using WMN technology

One of the shortcomings of WiFi is the short distance coverage in comparison to GSM for example. Interference associated with the congested nature of the 2.4 GHz frequency band is not much of an issue in rural areas due to the low density of devices operating in the same frequency band. However, when end devices are run in infrastructure mode (i.e. when devices need an Access Point to connect to the network), in order to provide connectivity across a large area, a dense deployment of Access Points is needed. Even though the IEEE 802.11S (an amendment to 802.11) standard defines the interconnection of wireless devices to form a mesh network, there is still several challenges that need to be addressed.

Firstly, Access Points in close proximity can coexist only if they communicate on orthogonal channels [54]. Interference occurs when two or more links in range of each other simultaneously operate on the same frequency. Solutions for this problem fall in two categories namely, scheduling approach and multi-frequency approach.

Scheduling approaches aim to schedule the nodes' transmissions such that no two nodes within range of each other on a multi-hop route simultaneously operate on the same frequency. This inevitably introduces delays especially as the node density increases. On the other hand multi-frequency approaches involve the use of nodes operating in more than one frequency. One way is to use a radio that supports multiple frequencies. The alternative is to use multiple radios each operating in a particular frequency. The advantages of using multiple radios is that it improves network capacity and reduces interference as well as latency however, it increases network complexity.

Besides the need for Access Points to coexist, the need to support mobile users brings about its own set of challenges in that uninterrupted connectivity should be ensured despite the high potential for user mobility. The difficult part is ensuring that there is no service disruption as users transition from one Access Point to the other.

The other challenge with 802.11 networks is that WiFi is line of sight communication. This implies that for communication to take place, nodes have to be placed within line

of sight. This does not necessarily mean that if you can see it then you can connect to it. In this context, line of sight refers to the electromagnetic line of sight which takes into account the needed Fresnel zone clearance. Line of sight can be hard to achieve in heavily rugged and undulating terrains. This challenge can be circumvented by using high towers or mounting nodes on terrain summits to overcome obstacles and provide Fresnel zone clearance needed for communication. The other solution is to use a non line of sight alternative to WiFi such as TV White Spaces (TVWS). However, TVWS technology is still in its infancy and hardware capable of taking advantage of the TVWS is still required.

## **2.6 Conclusion**

This chapter has provided a context and established the information needs of rural communities through a comprehensive literature survey. The chapter has highlighted some of the challenges rural communities are fraught with and includes a basis for an ideal communication and information dissemination system for users in these areas whose main occupation is small scale farming. The related work on phone based information systems for the rural population rely on the use of the cellular network. This dissertation proposes the use of VoIP over the community wireless mesh network infrastructure as a cost effective alternative to support the information system required by users.

## Chapter 3

# System design and implementation

This chapter discusses the design and implementation of ubiquitous mesh networking application to mobile communication and information dissemination. Bearing in mind the predominant occupation (small-scale farming) of the target users, the system draws from the drought early warning system, which comprises three major components, namely risk assessment, monitoring, and communication and dissemination of information as shown in figure 3.1.

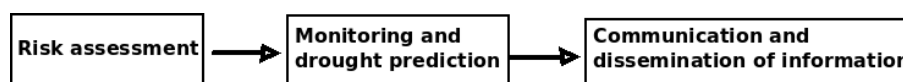


FIGURE 3.1: Components of a drought early warning system. Source: adapted from a framework for predicting droughts in developing countries using sensor networks and mobile phones [1].

Considering the components of a drought early warning system depicted above, the proposed system is centred on the information dissemination and communication component of the early warning system. The system utilises the existing wireless mesh network infrastructure to provide additional services. The first service that was considered is the use of the wireless mesh network infrastructure as a back-haul to islands of wireless sensor networks and having an application which allows users to query the network for information about sites being monitored (section 3.2.2). Another service that was considered is the implementation of Internet Protocol based radio for information dissemination (described in section 3.2.3.5). The other service considered is an extension of the existing primary service which is voice over Internet Protocol (VoIP) communication. The existing service implementation restricts the users to *hardphones*. Hardphone

is a term used to refer to the traditional wired telephone handsets. This dissertation considered the use of a *softphone* to facilitate community-wide mobile communication using the wireless mesh network infrastructure (section 3.2.3.6). A softphone is an application which runs on a computer/laptop or smartphone and allows users to make VoIP calls. Besides supporting mobility, using a softphone brings potential for added features such as real-time video streaming alongside the voice call. The other service that was added is a message board that allows users to post voice messages or questions on various subjects and also browse through others' posts (described in section 3.2.3.1). The implementation environment included Asterisk, VoiceXML, Speect TTS, icecast, Linux, Python, Lighttpd and CouchDB. The underlying hardware included, encased Alix system boards with its associated peripherals (described in section 4.1.1).

### 3.1 Design motivations

The system is targeted at a mainly small scale farming community in rural areas where limited textual literacy is expected. One challenge arising from that is the need to provide the information in a format that is suitable. The other challenge is to provide the service using technology that the user already owns and is proficient with. Among the options available is the use of visual aid to convey the information. However, that would restrict service access to users with a computer or high end smartphone. The second alternative is to deliver information in audio format. The second option is more favourable because it helps in meeting the objective of reaching a larger audience in that voice messages can be delivered even to commodity mobile phones. Using a voice service to disseminate information is a much more ubiquitous channel of access because of its availability on all mobile handsets. In addition, it has been found that textual interfaces are unusable by users with low textual literacy and difficult to use by novice users [55].

In order to overcome the challenge associated with providing information to users with low textual literacy, the design includes a translation gateway whose task is to translate textual information obtained from sensor readings or supplied by other system users into voice messages. Small scale farmers can then access the information by dialling a service number.

In addition, the targeted rural areas are characterised by a lack of a constant supply of electricity. This means that the system has to rely on low power devices that can harness available solar powering mechanisms.

## 3.2 System framework

The overall system architecture is depicted in figure 3.2. The various system components are described in the following sections.

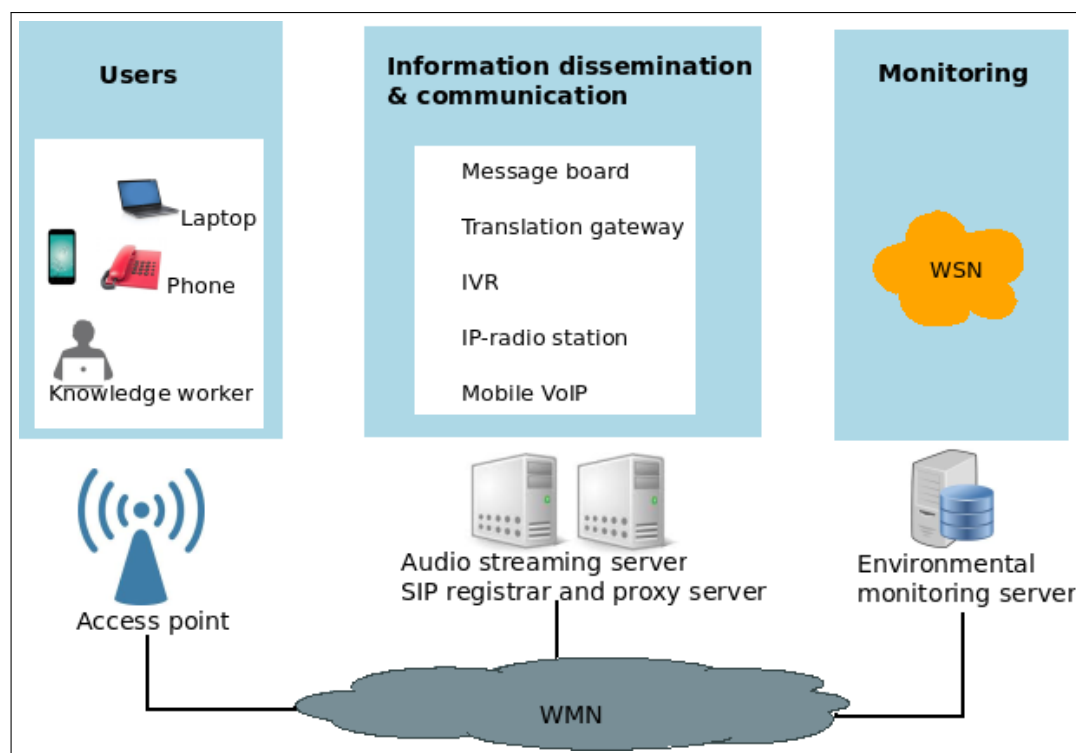


FIGURE 3.2: Mobile communication and information dissemination using wireless mesh network system architecture. Inspired by DFAS [2].

### 3.2.1 Users

Users include small scale farmers or members of the community as well as appropriate personnel from the agricultural office as an example. The term *knowledge worker* has been used to refer to the field staff from the agriculture department responsible for the area. The role of the knowledge worker is to analyse the sensor data, provide content for the community IP-based radio and respond to user queries.

The framework incorporates the handphones (also referred to as VoIP telephone handset) capable of connecting directly to the network. Analogue telephones may also be connected to the network using an Analogue Terminal Adaptor (ATA). The ATA has both RJ-11 and RJ-45 connectors and allows analogue telephones to connect to an IP network. A softphone running on mobile devices allows users to establish a SIP connection for VoIP.

---

The system allows users to dial a service number and navigate the system through voice prompts to access information. In addition, users are able to make phone calls to any other user registered on the network and have a voice conversation.

### 3.2.2 Monitoring

The deployment of sensor nodes consists of various types of sensors connected to the mote. These sensors respond to environmental stimuli and data is collected via the analogue to digital converter part of the mote's computational unit and the mote sends the collected environmental data to the sink node or base station. This data is then passed to the gateway connected to the sink node for additional processing. This multi-hop model allows the WSN to extend coverage over areas significantly larger than a single node's transmission range. In addition, nodes use less power because of the shorter transmission distance through this localised communication. Data that can be obtained from the WSN includes temperature, humidity, air pressure, precipitation and others depending on the environmental conditions of interest and the availability of specific sensors.

Data from each of the wireless sensor networks deployed in areas of interest is collected by the sink node and sent to the manager node as illustrated in figure 3.3. The designated manager node is ideally a computer with sufficient processing and storage capabilities but, due to power constraints in rural areas, the Alix system is used. The WSN can span a large area of phenomenon interest thereby collecting a large amount of data. The data can be accessed remotely by connecting to the community-wide wireless mesh network or via the Internet once a mesh network Internet gateway is put in place. Applications are required to aggregate and process the data across the network. Using data gathered from sensors processed in real-time is known to provide accurate information for purposes such as prediction of drought [43].

To aid in decision making, the data is processed with two objectives. Firstly, data from sensors is processed for purposes of assessing the prevailing conditions at the sites. The second objective is to infer the future conditions based on the evolved pattern. For instance, in the case of drought monitoring, which is considered as a situation where normal plant growth is negatively affected due to soil moisture deficiency resulting from a water shortage within a given period of time, rainfall is considered as the main source of water and temperature tends to affect the rate at which water is lost from the surface. Therefore, rainfall, soil moisture and temperature may be considered as key variables in monitoring drought.

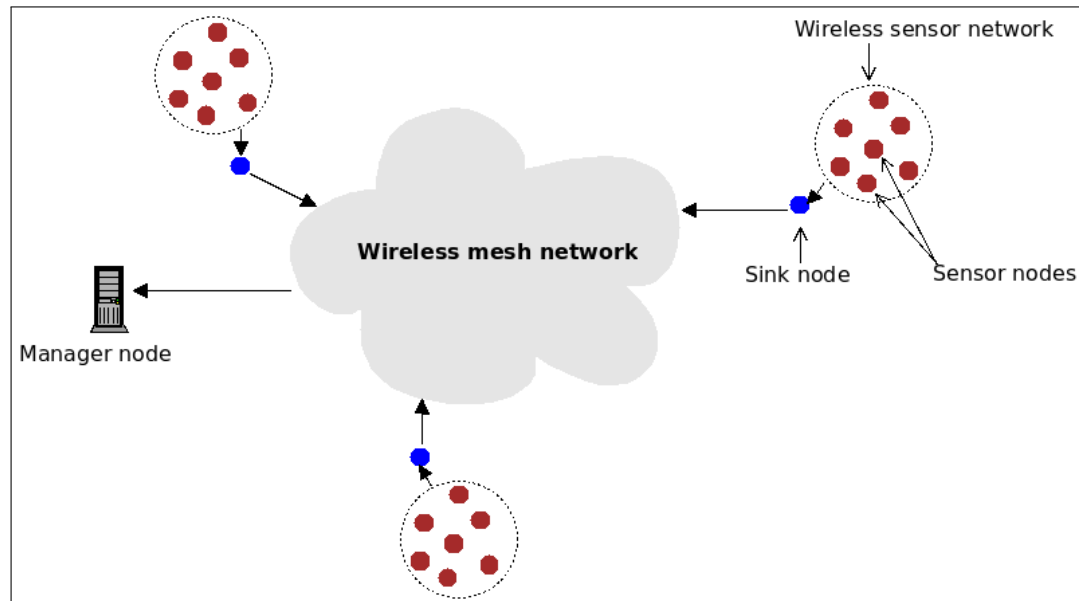


FIGURE 3.3: Using the wireless mesh network as back-haul to islands of wireless sensor networks.

Studies [56], [57], [58] have shown that wireless sensor network performance is influenced by the sensor node placement pattern. The studies show that regular formations such as circular, square or equilateral triangle cell shapes have advantages such as reduced packet flooding which in turn saves the bandwidth of the network and saves power thereby increasing the lifespan of the network because the batteries powering the nodes tend to last longer. Therefore, while the goal is to monitor the entire area of interest, these sensor nodes need to be systematically placed.

### 3.2.2.1 System requirements

The following were identified as services which the system should provide:

- Allow users to view WSN data via web interface
- Allow users to dial a number and receive information
- Send alert (text message or voice call with recorded message) when defined thresholds are reached.

In this design, the sink nodes or the WSN gateways deployed in the field send the collected data to the designated manager node as illustrated previously in figure 3.3. The main purpose of the manager node is to process the collected sensor data and have the information useful for decision making purposes. For instance, when soil moisture



---

falls below a given threshold, the farmer is alerted about the situation via text message or automated voice call with a pre-recorded voice message. The farmer can then decide to turn on the irrigation pumps. Conversely, when the soil moisture rises above a set threshold, the farmer is again alerted and can make an appropriate decision such as switching off the irrigation pumps. By so doing, the system helps the farmer to utilise the resources optimally. Ideally, WSN technology includes actuators to automate the actions but, this dissertation does not cater for actuators.

In some other implementations, the sink node sends the collected data to a central station via General Packet Radio Service (GPRS). One disadvantage of using GPRS is the cost associated with the service. At the time of writing, the cost of GPRS was about R2/Mb (about 0.20 US Dollar). In this design the WMN is used to link the WSN with the central data collection point. One advantage of WMN is that a range of wireless radio technologies such as 802.11 and 802.15 (i.e. WiFi, Bluetooth and Zigbee) can be supported concurrently. WMN offers the flexibility needed to connect WSN deployments to data processing units and allows remote access to devices or nodes.

The alert mechanism described herein was implemented using Asterisk's *call file* and *Asterisk Gateway Interface* (AGI) features. The manager node interacts with the PBX via AGI. The AGI allows the implementation of call control externally to the PBX. The call file feature allows the automation of communication channel opening. Once this channel is automatically open, it can be used to send a text message to a client or make a voice call with a system generated voice message.

### 3.2.3 Information dissemination and communication

The system uses Asterisk, a software PBX system that is built around the concept of modules where a module is defined as a component loadable at runtime to provide specific functionality [37]. Of particular interest to this dissertation was the application modules which are used to specify what actions to take for a given call. When a call reaches the server, the call handler takes one of two actions:

- Service the query via IVR (explained in section 3.2.3.3) when a service number is dialled; or
- Relay the call to establish connection between devices when other client numbers are dialled.

Making a device to work with Asterisk is a two part process in that firstly, the device needs to be known by Asterisk and secondly, the device needs to know about Asterisk.

This is achieved by defining SIP clients at the Asterisk server and allowing the clients to register with the Asterisk SIP server as illustrated in section 3.2.3.6, figure 3.9.

### 3.2.3.1 Voice-to-voice, text-to-voice community information system

From the background chapter it has been pointed out that language offering and input modality are among the main design concerns of a voice-based information system. In order to make the information relevant for the people as well as handle the language concerns, it is best to let the users create the content for themselves. The designed system framework relies on the inclusion of an operator for platform regulation and supply of specialised information. The following are some of the key service features:

- Users need to be able to post messages, questions or comments.
- Users need to be able to browse other users' posts.
- Users need to be able to comment on other users' posts.
- The designated Operator needs to be able to upload information as voice or as text via a user friendly interface. When users dial the service number the information entered as text is accessible as voice via the phone as illustrated in figure 3.4. The mechanism employed in the conversion of text to voice is explained in section 3.2.3.4.

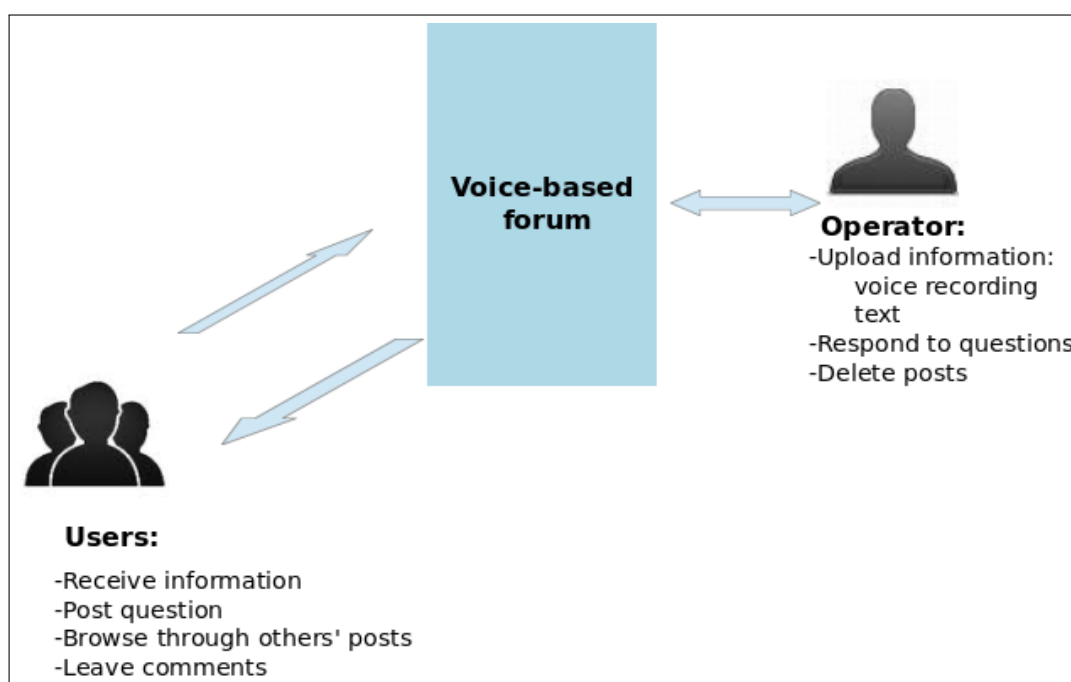


FIGURE 3.4: Rural community information system.

The voice-based forum shown in figure 3.4 was designed to meet the identified service requirements. It is an asynchronous voice-based forum aimed at enhancing information sharing. The system was implemented by repurposing Asterisk's call recording and voicemail features to meet the objective. The basic concept is to have the calls coming to the service number recorded similar to how voicemail works. The voice messages are then made accessible to all entrants of the forum. This provides a cost effective asynchronous voice chat system that is cost effective in terms of the monetary cost associated with platform acquisition as well as the hardware requirements of the system.

**Regular user usage scenario:**

*Dial service number*

*Select information category*

*Select task:*

*I post a message*

*II browse through messages*

*III exit*

*Case I:*

*leave message*

*return to menu*

*Case II:*

*Loop*

*Listen to messages*

*Do something: comment/next*

*Return to menu*

*Case III:*

*End*

**Operator usage scenario:**

*Dial service number*

*Select category*

*Select task:*

*I Leave a message*

*II Browse through messages*

*III Exit*

*Case I:*

*Leave message*

*Return to menu*

*Case II:*

*Select category*

*Loop*

*Listen to message*

*Do something: delete/comment/next*

*Return to menu*

*case III:*

*End*

**3.2.3.2 Organisation of information**

The organisation of information was influenced by the areas or categories of information needed, namely education, government, agriculture, health, business, people, community news, and general information (*see section 2.1*), which in turn dictated the structure of the menu. One might suggest merging some of the categories but, having a granular organization of information makes it easy for someone to find the information that they are looking for. The information to be stored is either going to be static, quasi-static or dynamic. Static information is information which does not change frequently. On the other hand, dynamic information such as community announcements changes much more frequently. The information has to be organised such that it is easily accessible.

---

The initial thought on the choice of organising the information was to use a hierarchical structure. This approach is based on Shneiderman's principle of visual information seeking mantra: overview first, Zoom and filter, then details on demand [59]. However, it was discovered from literature [60] that while Shneiderman's mantra of information seeking works well for graphical user interfaces (GUI), it may not be ideal for voice-based user interfaces. This is because such an approach results in a voice-based user interface that is tasking on the part of the user. Visual search and auditory search happen to differ fundamentally such that theories of visual search may not be applied to the auditory domain.

### 3.2.3.3 Interactive Voice Response (IVR) design considerations

IVR is a telephony interface characterised by voice prompts in response to a call for purposes of automated information or service provision via the telephone. Users may provide input using voice however, the designed system does not employ speech recognition. This is because of the inherent limitations of speech recognition such as the broad sources of variability which tend to decrease the accuracy of such systems. Instead, DTMF (Dual Tone Multiple Frequency) or telephone keypad touch-tone is used for interaction with the server.

One of the major things to consider when designing an IVR is the language offering but, due to time and resource constraints, the target language for the designed system was restricted to English. In addition, voice based telephone menus have traditionally consisted of a 0 (zero) option to reach a live person and so, the option to dial 0 in the IVR to reach a live person should be included where applicable. Furthermore, research has shown that people tend to not pay full attention to the voice prompts until they hear the choice that is of interest to them [37]. Therefore, the main menu should be designed in such a way that the users are given a description of the menu option before giving them the digit to press for that option. For instance, instead of saying "*press 2 for today's temperature*", the system should say "*for today's temperature, press 2*". Moreover, users may not pay attention throughout the entire menu and so, it is essential that they be provided with the option of listening to the choices again. In addition, users are bound to make mistakes and so, there is need to gracefully handle errors and notify the users whenever an invalid option is chosen.

Given an increasing number of menu options, there are two main possible general approaches. One solution is to have a long menu inclusive of all options while the other approach is to have multiple shorter menus with sub-menus. From the literature consulted [60], there seems to be two opposing camps on which approach users accept that

allows them to complete their tasks effectively. The general arguments are that because of the human memory constraint, the menu should be kept short or else users will not be able to remember the options. When there is a large number of menu options, they need to be broken up into shorter menus and then include sub-menus. However, other researchers have argued that nested menus tend to be confusing for users and that a longer flat menu is preferred because it is more efficient and effective.

There are many factors which affect the user experience in addition to the menu length such as the users' cognitive ability that need to be put into account when determining the suitable approach. Therefore, the ultimate solution might lie somewhere midway between the two opposing schools of thought.

Figure 3.5 shows the adopted IVR design model proposed by Grover et al [61] for textually low literate users in the developing world. The term 'developing world' has been used to refer to populations that have low textual literacy and are technologically inexperienced.

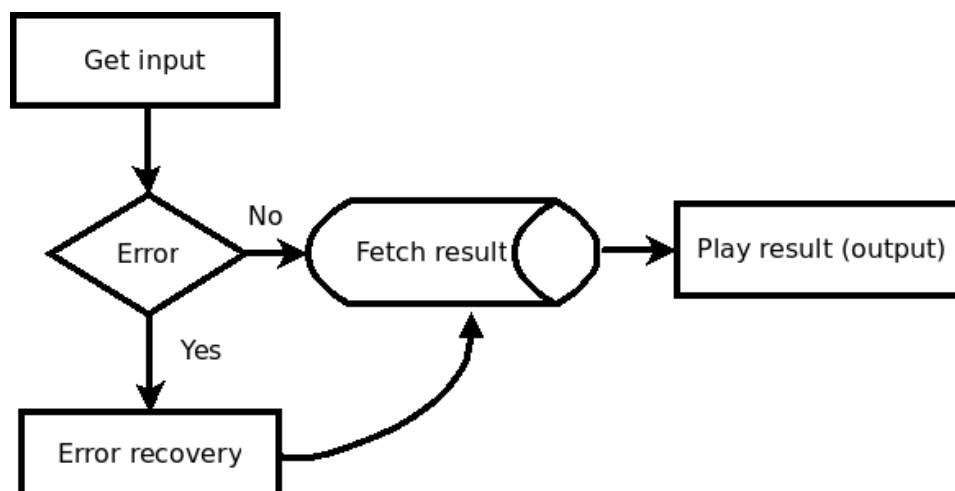


FIGURE 3.5: IVR design model.

The model consists of three components namely get input, error-recovery and output. *Get input* refers to the mechanism with which the user interacts with the system. In this design the menu options are presented in numerical form, which can be selected by pressing on the phone keypad. *Error-recovery* is the component that validates the user input. In the event of an error in the user input, the user is guided into making a correct request and confirmation prompts are used as a way of recovering from errors. *Output* refers to the means of presenting the information that the user requested.

### 3.2.3.4 Translation gateway

A translation gateway is an enabler that has the potential to enhance access to network resources by catering for users with no browser or Internet connectivity. The translation gateway enables this by integrating speech and telephony into the network resource access application. Using the client-server paradigm, users interact with a textual data storage mechanism from a telephony application as shown in figure 3.6.

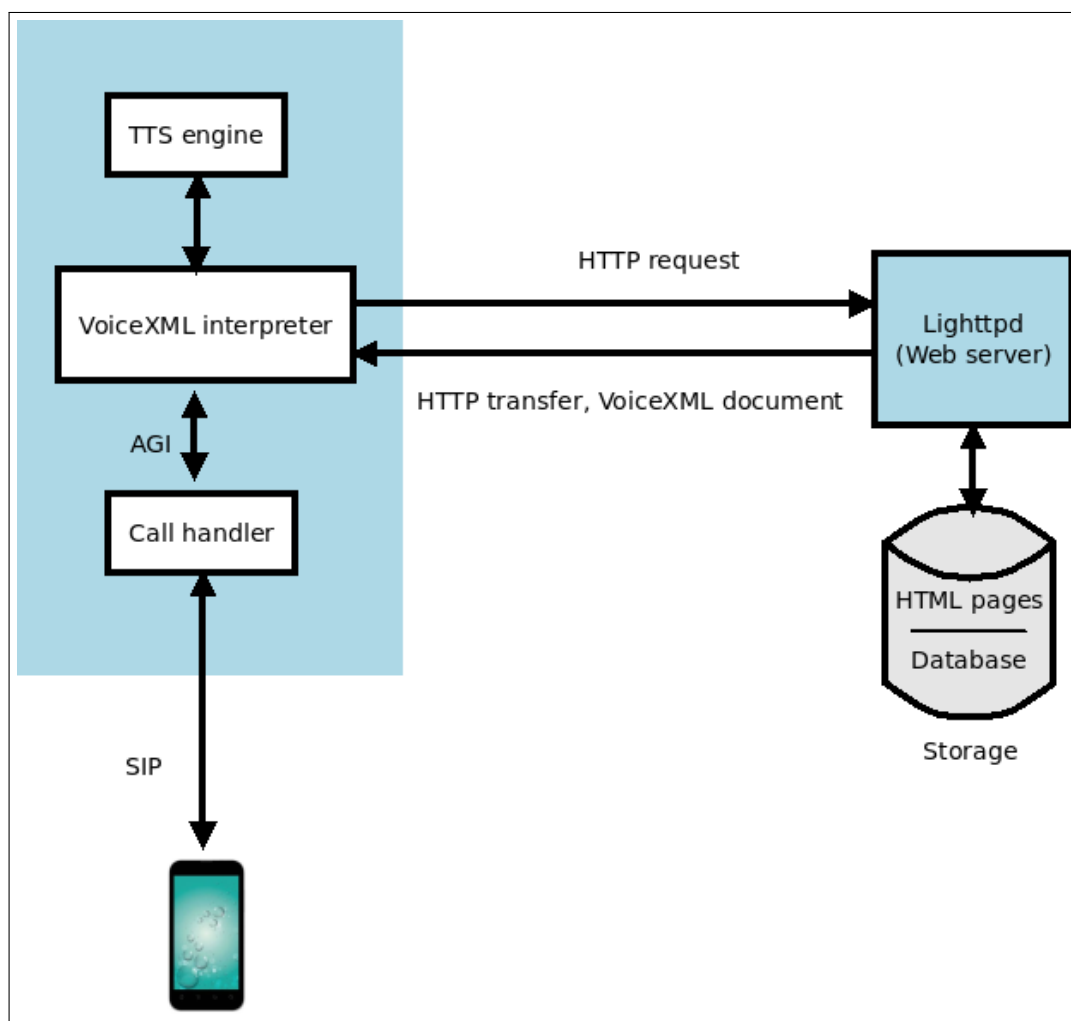


FIGURE 3.6: High level view of the translation gateway.

The translation gateway consists of the following three major components:

- *Call handler.* The call handler implemented in Asterisk PBX determines what to do with the call depending on the extension dialed. If the extension is that of another SIP client, the call handler instantiates a connection with the client thereby establishing a SIP connection between two clients for voice call purposes.

- *VoiceXML interpreter*. When a user dials a service number, the call handler maps the phone number to a URL. The gateway interacts with the web server using HTTP. When the gateway makes an HTTP request to the web server, the web server responds with a VoiceXML document. The VoiceXML document is very similar to an HTML document but contains voice dialogues and tags used to determine which parts of the text to synthesise to speech.
- *Text-to-speech engine*. The text-to-speech (TTS) engine automatically converts text to speech.

The web server was implemented using *Lighttpd*<sup>1</sup>. Lighttpd was deemed best fit for low-power devices because of its low memory footprint and CPU requirements in comparison with other web servers. The web server was considered as the appropriate way to organise static and quasi-static information which needs to be accessible via the phone or the browser.

The translation gateway uses text-to-speech technology to convert textual information to equivalent audio form so that it can be accessible via the telephone. Even though the current text-to-speech technology cannot entirely replace professionally recorded information, using recorded prompts and text-to-speech tools to automate the process of translation is useful in applications consisting of dynamic information that has to be communicated to users using telephony.

### Translation process

When the user through the IVR (described in section 3.2.3.3) requests for information which is in text form, the call handler interacts with the VoiceXML interpreter which in turn interacts with the TTS engine via *Asterisk Gateway Interface* (AGI) to provide textual information in audio form as illustrated in figures 3.6 and 3.7. The AGI facilitates the interaction of Asterisk with external processes.



FIGURE 3.7: Text to voice translation process.

The TTS engine does not support the streaming of output in real-time therefore, the resulting audio has to be buffered before being streamed to the user. Furthermore, for use with Asterisk, WAV files encoded in 16-bit, 8000 Hz, mono format are recommended

<sup>1</sup><http://www.lighttpd.net/>



[37]. This audio file format is easily handled by Asterisk and can easily be converted to other formats supported by phones without distortion or loss of quality.

For test purposes, the system was implemented using *Speect* TTS engine [62], available as open source for text-to-speech conversion. By using such readily available tools, the study time is reduced as concentration is shifted towards the dissertation objectives. However, where natural voice quality is required, commercially available TTS engines should be used for improved speech quality.

### 3.2.3.5 IP-based community radio

There are two approaches to the implementation of Internet radio. One approach is where audio files are made available for download and listeners can play them in their own time. The other approach is where audio is streamed to the listeners in real time using pre-recorded audio files or live audio feed.

Internet radio service is ideally available to anyone with Internet access. This dissertation considered the concept of Internet radio as an added mechanism for information dissemination in a rural community using a wireless mesh network, hence the use of the term, IP-based community radio. In order to listen to the IP-based radio, the user simply enters the streaming server's URL or IP address in the stream player. Implementing IP-based radio on the community network comes with the advantages proffered by WiFi namely, no broadcasting license required.

The high level architectural view of the Internet radio consists of two major components namely, broadcast or streaming server and encoder or source client. The source client encodes the audio input and feeds it into the streaming server. The source client can either read audio files from storage or encode live audio coming in through the sound card. The streaming server then streams the audio to client listeners over standard HTTP as illustrated in figure 3.8.

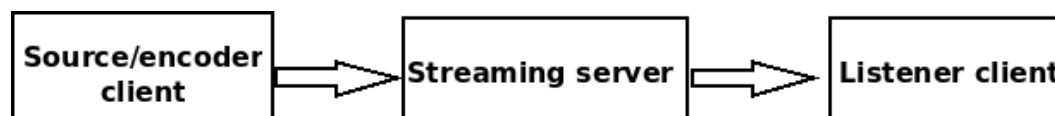


FIGURE 3.8: Community IP radio station audio streaming overview.

The open source tools Icecast <sup>2</sup> and Ices <sup>3</sup> were used for the streaming server and source client respectively. Icecast supports mp3 or Ogg Vorbis audio file formats. The streaming server and the source client can run on the same machine but, in our implementation

<sup>2</sup><http://www.icecast.org/docs.php>

<sup>3</sup><http://www.icecast.org/docs/ices-2.0.2/>

---

we ran the source client and streaming server on separate devices to split the processing load. The listener client or decoder client can be any stream player running on the end user device.

Internet radio content can be a predefined list of audio files or live feed i.e. a person speaking through the sound input mechanism in real-time. It is possible to stream any type of audio including synthesised text to speech. As a technical requirement, the audio file only needs to be in one of the formats supported by the streaming server. But, in order to increase the likelihood of a voice-based service being adopted by the target community, researchers recommend finding a local partner organisation within the user community to act as intermediary thereby ensuring that the content is kept relevant [63].

### 3.2.3.6 Mobile communication using wireless mesh network

The wireless mesh network has been used to provide telephone communication in rural communities using Voice over Internet Protocol (VoIP). These communities use wired phones (sometimes referred to as *hardphone*) which connect directly to the network or via an analogue terminal adaptor. The purpose of the analogue terminal adaptor is to allow analogue telephones to connect to the network. As pointed out earlier, these phones are fixed in one place. The aim of this dissertation was to investigate how to enable VoIP communication for mobile phones using the wireless mesh network infrastructure. Adding mobile VoIP support contributes towards ensuring that all telephone services available on the network are made accessible via the mobile phone as well.

To add VoIP support for mobile phones, the use of a softphone was considered. A softphone is an application that runs on a smartphone to establish a VoIP connection. For the VoIP gateway and private branch exchange (PBX), Asterisk was used. Asterisk supports several protocols notably, H.323, Media Gateway Control Protocol (MGCP) and IAX. However, the current Mesh potato being used in rural communities consists of analogue telephony adaptors which support the Session Initiation Protocol (SIP). In addition, a comparative study of VoIP protocols with Asterisk shows that SIP has relatively lower processor and memory utilization [64], therefore, for easy integration of mobile support, SIP was the VoIP protocol of choice.

Even though SIP is a peer-to-peer protocol, centralising the signalling process simplifies the management tasks and is known to provide better performance [52]. Once signalling is complete, the SIP server is transparent and the two SIP clients communicate directly as illustrated in figure 3.9. But in essence, a connection is composed of two calls: firstly, a SIP client calls the SIP server then, the SIP server calls the other SIP client.

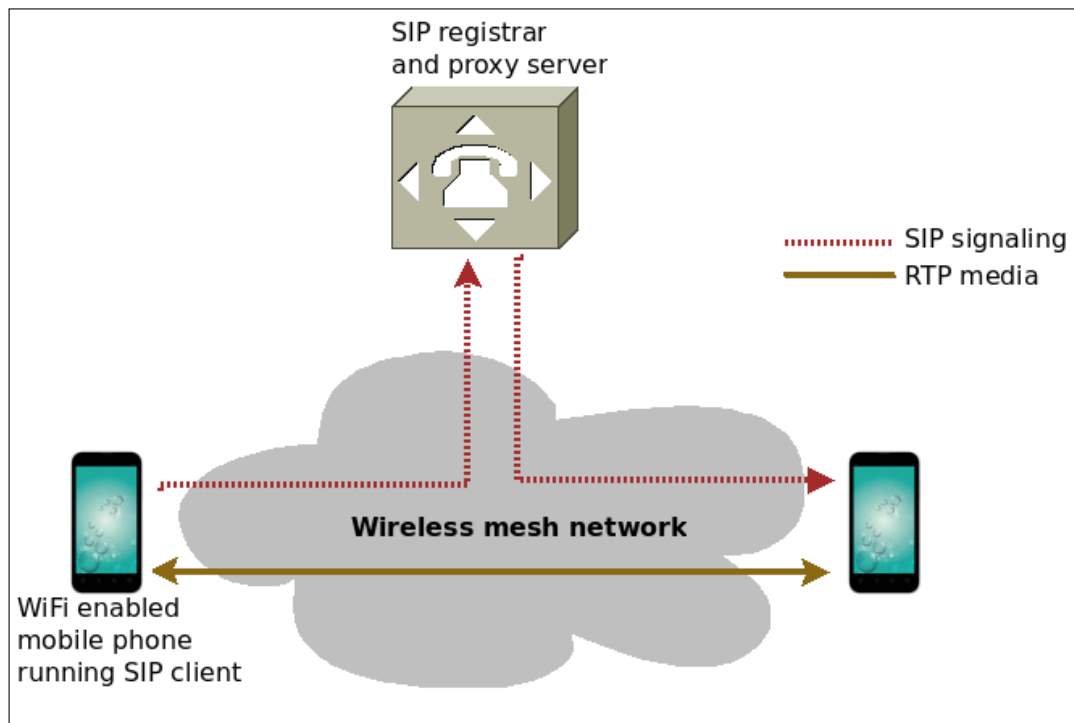


FIGURE 3.9: Using wireless mesh network for mobile communication. Source: adapted from the SIP trapezoid [3].

SIP does not transport media between endpoints, once a SIP connection is established, media (i.e. the actual voice) is transported between the two endpoints using *Real-time Transport Protocol* (RTP) which in turn is encapsulated in UDP. Studies have shown that using connectionless UDP for VoIP provides better quality of service compared to the connection oriented TCP [65].

There has been some *infrastructure-less mode* VoIP implementations such as the *Serval project* [51]. Infrastructure-less mode is where end user mobile devices alone make up the network infrastructure. These mobile devices connect directly to other mobile devices within range thereby creating a network on an ad-hoc basis with each device taking part in the relaying of voice traffic on the network. In order to achieve this, devices are operated in multi-mode i.e. a device allows other devices to connect to it but at the same time is capable of connecting to other devices and act as a router by forwarding network traffic. The infrastructure-less VoIP has a number of advantages such as non reliance on infrastructure, which requires some capital investment and also poses a potential point of failure. However the infrastructure-less approach is still in its infancy and more work is needed before users can have a reliable service. Furthermore, the infrastructure-less implementation requires a high density of mobile devices over the coverage area in order to provide sufficient end to end connectivity. The rural communities considered in this dissertation are sparsely populated and so infrastructure is required for complete connectivity among households.

### 3.3 Modifications needed on the WMN infrastructure to support new services

A voice-based service such as what has been proposed requires uninterrupted connectivity. The network should be transparently accessible from all 802.11 devices. However, users accessing the network via mobile devices are inevitably going to roam around and they need to do so within the mesh network coverage area without experiencing any connectivity disruptions. One challenge arising from mobile-based VoIP is the need for *hand-off* support for roaming users. For instance, what happens when a user makes a call to another user using SIP and then moves from one Access Point to another whilst in the middle of the conversation?

Hand-off algorithms and mechanisms are well implemented in cellular networks however, from the literature surveyed at the time of writing, there does not seem to be a standard roaming specification for 802.11 based wireless networks.

#### 3.3.1 Supporting mobile users

The community based wireless mesh network under consideration was originally designed for static VoIP communication i.e. users use hardphones fixed in one place to make calls. Using the infrastructure to support mobile users poses a different set of challenges. To start with, the current VoIP protocol implementation requires that the hosts engaged in a session maintain the same IP address through the duration of the session. However, with mobile devices, the mobility might result in hosts changing the IP address as they switch from one Access Point to the other. As pointed out in the previous section, the 802.11 standard does not define a roaming specification. In addition, the ideal ubiquitous wireless mesh network should make no assumptions about any special network driver or software running on the client. All network connectivity is solely reliant on the Media Access Control (MAC) and Internet Protocol (IP). Therefore, in order to add support for mobile users, the hand-off approach proposed by Musaloiu-Elefteri [66] was adopted. The key steps towards achieving uninterrupted seamless connectivity for mobile users are:

- ensuring that clients have the same IP information
- ensuring that clients have the same gateway information
- ensuring that more than one Access Point handles the moving client.

---

Using *Dynamic Host Configuration Protocol* (DHCP) running on mesh nodes to configure client IP information, an association is made between the client's MAC address and IP address. Thus when a client requests an IP address, the DHCP server computes a hash function on the client device's MAC address to derive the IP address. This ensures that clients obtain the same IP address and network mask from anywhere across the wireless mesh network coverage area.

Furthermore, the client sets the same default gateway address regardless of which Access Point they are connected to. This gateway address is supplied by the DHCP server and in reality is not used by any node in the network hence it is referred to as a *virtual gateway* address. Clients are able to access the network via this virtual gateway by associating the virtual IP address with the mesh nodes' MAC address. This is achieved through the use of gratuitous *Address Resolution Protocol* (ARP)<sup>4</sup> packets sent by the mesh nodes. In slight deviation from how ARP works, gratuitous ARP packets are not sent in response to an ARP request, instead they are periodically broadcast to the network groundlessly which allows network hosts to update their ARP tables [67].

In addition, allowing the Access Points to monitor and share clients' connectivity quality information lets the Access Points coordinate the transition of the client from one Access Point to another. Once these mechanisms are put in place, it seems to the mobile client as though they are constantly connected to the same Access Point despite the mobility. This approach is aimed at eliminating service interruption during the hand-off by reducing the amount of time it might take a client to reconfigure its IP after transitioning to a new AP.

### 3.3.2 Optimizing performance

Running applications on the network for various additional services is bound to increase the amount of traffic and resource consumption, which might in turn affect network performance. It may be said that increasing the services offered by the network and improving the quality of the services are two opposing goals. The overall performance of a wireless mesh network depends in part on the performance of the mesh nodes or routers. But these nodes have limited CPU and memory capacities. This is even more so in rural communities with no electricity where you have to rely on low-power devices.

Musaloiu-Elefteri [66] opines that routing all the traffic through user-space is convenient but strains the router's resources. Particularly, user-space routing significantly increases CPU utilization. Therefore, in instances where router processing capability is a concern,

---

<sup>4</sup>Address Resolution Protocol is used to map a network layer protocol address such as an IP address to a data link layer MAC address

kernel-space routing should be employed as far as possible in order to optimize performance. The increased CPU overhead associated with user-space routing is attributed to increased memory copies and context switches employed in the user-space routing mechanism.

### **3.4 Conclusion**

This dissertation considered enhancing communication and information access amid the constraints of a rural context. The designed mobile communication and information dissemination framework provided the required basis upon which to integrate the sub-systems and implement the system. In the next section, the system is tested in a laboratory setting and the results presented.

## Chapter 4

# Performance evaluation

This chapter describes the tests which were carried out to evaluate the addition of services namely mobile VoIP communication, IP-based radio station and the text-to-voice translation gateway on the wireless mesh network.

Evaluating network service requires quantifying and analysing the quality of experience (QoE) for purposes of evaluating and managing the network services. The existing methods of service evaluation can be placed into two categories: subjective and objective evaluation. Subjective methods of evaluating the quality of experience (QoE) are based on user scores through surveys. For this dissertation, the evaluation of the services was done by carrying out quality of service (QoS) measurements and then using *QoS-QoE* correlation models [68].

This method tries to relate the objective network service conditions with the human perception of the quality of service. While there might be differences between the measured QoS and the QoE perceived by the user, for this dissertation, the measured quality of service is considered to be sufficiently closely related to the quality of experience for the end users. Different models for objective evaluation known to correlate objective measurements with subjective scores have been used for each of the services. The choice of model to use described herein is based on the appropriateness of the method, bearing in mind the parameters that each of the models puts into account in order to come up with the best estimate of QoE.

The next sections describe the WMN test-bed and a set of experimental measurements conducted to investigate the following:

- i. the mobile VoIP call quality
- ii. the VoIP capacity of the underlying hardware

- iii. the quality of the audio and the impact of audio streaming on the network
- iv. the quality of speech generated by the text-to-speech translation gateway.

The purpose of these tests and measurements is to ascertain the implementation of such services on low-power devices namely, the *Alix system* (described in section 4.1.1) and also to assess the quality of service rendered. The introductory chapter of this dissertation laid down the motivation behind the proposed services. The performance results obtained through the tests and experiments described here will help in determining the technical feasibility of service implementation using low-power hardware devices bearing in mind the computation and storage limitations of such devices.

## 4.1 WMN test-bed design

A number of steps were taken in building the various parts of the WMN test-bed. To start with, suitable hardware had to be sourced and assembled. Then decisions had to be made on the suitable operating system software to be installed. Additionally, programs essential to run the WMN had to be installed. Installation of one component often led to the need to install another in order to meet the library dependency requirements. Lastly, the WMN test-bed itself had to be tested to ensure that it was working as required before conducting any experiments.

### 4.1.1 Node description

The nodes used in the experiments consisted of Alix 2D2 system boards shown in figure 4.1, from PC Engine <sup>1</sup>. The Alix 2D2 consists of 500 MHz AMD Geode LX800 CPU and 256 MB DDR DRAM.

Each of the nodes was equipped with a miniPCI cm9 802.11 a/b/g Atheros wireless card. Support for different wireless modes is a desired feature of the wireless card because it allows for a wider range of test scenarios to be carried out without needing to change the hardware. Furthermore, the nodes were fitted with an RP SMA male antenna connected to the wireless card using the RP SMA female pigtail connectors. For internal storage purposes, a 4 GB Compact Flash card was installed onto the boards.

---

<sup>1</sup> <http://www.pcengines.ch/>





(a) Alix board.



(b) Alix board in a case.

FIGURE 4.1: Alix system.

The light weight version of Linux known as *Debian for Alix* was installed on the nodes and network interfaces had to be configured to achieve the required connectivity. BATMAN advanced [69] was installed and configured on the nodes for the multi-hop forwarding of packets on the WMN test-bed. For test scenarios where clients had to connect to the WMN, *hostapd* was used. Hostapd is a Linux wireless driver which lets the clients see the node as an Access Point (AP).

### 4.1.2 Enabling multi-hopping on WMN test-bed

From the observation made on the test-bed implementation, the transmission distance covered by the node is determined by its physical radio characteristics and the signal power output software settings. When a signal is transmitted from a node, it has an output power which is lost as it propagates through the medium. Multi-hopping can be enabled by either placing the nodes at distances sufficiently long or by using attenuators to attenuate the signal so as to force multi-hopping at given distances. The effective distances in the latter case depends on the attenuator specification. In order to enable multi-hop on the indoor test-bed implementation with limited space for node placement, it was discovered that removing the antenna, leaving only the pigtail provided sufficient signal attenuation to force multi-hopping at short distances. Enabling multi-hopping was essential in some of the experiments (for instance, section 4.4) conducted in order to mimic the real wireless mesh network characteristics. By varying the distances among the nodes, we were able to achieve desired physical network topologies needed for experiments such as the multi-hop single-path and multi-hop multi-path topologies illustrated in figures 4.3 and 4.4.

## 4.2 Determining the mobile VoIP call quality

In order to quantify the performance of the mobile VoIP solution, key parameters namely call setup time, maximum VoIP latency, delay, jitter and VoIP packet loss for voice calls have to be evaluated.

ITU-T provides one objective and one subjective method of measuring the quality of voice based communication. The objective method is referred to as the E-model whereas the subjective method is known as mean opinion score (MOS) [4]. The E-model is a mathematical model that combines factors which affects voice quality to quantify the quality of the call from sender to receiver as shown in equation 4.2. On the other hand, MOS relies on users to rate the quality of service on a given scale. Objective measurements have the advantage of being repeatable and can be carried out in real-time.

Objective quality measure may be expressed generically as:

$$Q = F(x_1, \dots, x_n) \quad (4.1)$$

where  $x_i, \dots, x_n$  are source and network parameters. Using the E-model, the output is a Transmission Rating Factor,  $R$  used to estimate the user satisfaction of the handset conversation from the listener standpoint. The value of  $R$  is given by the equation:

$$R = R_0 - I_s - I_d - I_e + A \quad (4.2)$$

where,

- $R_0$ : the signal-to-noise ratio
- $I_s$ : sum of real-time speech transmission impairments
- $I_d$ : sum of delayed impairments relative to the speech signal
- $I_e$ : equipment impairment factor such as the bit-rate for the codec
- $A$ : advantage factor.

The value of  $R$  can be converted to subjective measures such as MOS often used to express the quality of a voice call. The value of  $R$  ranges from 0 to 100 whereas the MOS score ranges from 1 to 5 with higher values indicating higher quality of service as shown in table 4.1. MOS values higher than 3 are considered acceptable whereas a score of 1 is said to be impossible to communicate. Alternatively, the E-model objective

<b>R</b>	<b>Quality category</b>	<b>User satisfaction</b>	<b>MOS</b>	<b>%GOB</b>	<b>%POW</b>
90 - 100	best	very satisfied	4.3 - 5	97.0 - 100	0.2 - 0
80 - 89	high	satisfied	4.0 - 4.2	89.5 - 96.9	1.4 - 0.3
70 - 79	medium	some users dissatisfied	3.6 - 3.9	73.6 - 89.4	5.9 - 1.5
60 - 69	low	many users dissatisfied	3.1 - 3.5	50.1 - 73.5	17.4 - 6.0
50 - 59	poor	nearly all users dissatisfied	2.6 - 3.0	26.6 - 50.0	37.7 - 17.5
0 - 49	unusable	not recommended	1.0 - 2.5	0 - 26.5	99.8 - 37.8

TABLE 4.1: E-model output scales and definition of categories of speech transmission quality. Source: ITU-T Recommendation G.107 [4]

results can be mapped to other subjective values such as Good or Better (GoB) or Poor or Worse (PoW) expressed in percentages as shown in table 4.1.

As pointed out earlier, subjective methods such as MOS do not provide for measurements to be carried out in real-time. Therefore, for this dissertation, estimation of QoS was done using an improved simplified version of the E-model [70]. This method tailored to

VoIP quality estimation focuses on important factors only thereby reducing some of the complexity with which the E-model is known for. Using the improved simplified version of the E-model the rating factor  $R$  is calculated as shown in equation 4.3.

$$R = R_y - I_d + A \quad (4.3)$$

Where  $R_y$  is a second order function,  $I_d$  is the average delay in a given period of time and  $A$  is the advantage factor associated with the communication system. ITU-T recommendation G.107 provides example advantage factor values for various types of networks as shown in table 4.2. The other parameters ( $R_y, I_d$ ) are calculated as follows:

$$R_y = aR_x^2 + bR_x + c \quad (4.4)$$

where  $a, b$  and  $c$  are codec specific coefficients shown in table 4.3 and  $R_x$  is given by:

$$R = R_0 - I_{codec} - I_{packetloss} \quad (4.5)$$

where,  $R_0$  is the signal-to-noise ratio. This encompasses noise from the circuitry as well as the physical surrounding. Because of the inherent difficulty in calculating the signal-to-noise ratio, ITU-T G.113 recommendation provides maximum theoretical value of 94.2 with no impairment or 93.2 with impairment which can be used for estimation purposes.  $I_{codec}$  is the codec factor, which represents the voice distortion arising from signal conversion. The  $I_{codec}$  value depends on the codec and can be obtained from the ITU-T recommendation G.113 as shown in table 4.4.  $I_{packetloss}$  is the percentage packet loss i.e. the percentage of packets sent by the sender but not received by the receiver in a given period of time.  $I_{packetloss}$  may be expressed as:

$$I_{packetloss} = \left(1 - \frac{N}{DS}\right) \times 100 \quad (4.6)$$

where given  $N$  packets,  $DS$  is the difference between the smallest and the largest sequence numbers. These sequence numbers can be obtained from sequence number field of the RTP header packets.

From equation 4.3,  $I_d$  refers to the one way delay i.e. the time it takes for the data from the sender to reach the receiver. For VoIP, this is the time taken for a voice signal from the sender's mouth to reach the receiver's ear. Owing to the factors which affect the

Communication system	Maximum value of advantage factor, $A$
Wired network	0
Mobility by cellular networks in a building	5
Mobility in a geographical area or moving in a vehicle	10
Access to hard-to-reach locations, e.g., via multi-hop satellite connections	20

TABLE 4.2: Advantage factor ( $A$ ) example values.

Codec	$a$	$b$	$c$
G.711	0.18	-27.90	1126.62
G.723.1 5.3k	0.039	-4.2	166.61
G.726 24k	0.046	-4.53	168.09
G.729A	0.063	-8.08	311.72

TABLE 4.3: Coefficients for the respective codecs

Encoder	Reference	Bitrate (Kbit/s)	$I_{codec}$ value
PCM	G.711	64	0
ACELP	G.723.1	5.3	19
ADPCM	G.726	24	25
CS-ACELP	8	-8.08	11

TABLE 4.4: Codec factor values.

time taken for data to get across to the other end of the network, the total delay may be expressed as follows:

$$I_d = t_i + t_{i+1} \dots + t_{i+n} \quad (4.7)$$

where  $t_i$  are the main elements contributing to the total delay namely, *transport delay*, *propagation delay*, *jitter buffer delay* and *Encoding/decoding/packetisation delay*. Encoding/decoding/packetisation delay refers to the delay introduced on the sender end as a result of the time required to accumulate voice input into frames and insert the frames into packets and transfer the packets to the transport layer, and the time needed at the receiving end to decompress the frames into voice samples.

Each of the measurements required to calculate the call quality were averaged over four runs.

$$I_{packetloss} = 1, I_{delay} = 1.2355$$

$$R = aR_x^2 + bR_x + c - I_{delay} + A$$

$$R_x = R_0 - I_{codec} - I_{packetloss}$$

$$= 93.2 - 1$$

$$= 92.2$$

$$R = 0.18 \times 92.2^2 - 27.90 \times 92.2 + 1126.62 - 1.2355 + 5$$

$$= 88.1557$$

Mapping the rating factor  $R$  to MOS score:

For  $0 < R < 100$

$$MOS = 1 + 0.035R + R(R - 60)(100 - R)7.10^{-6}$$

$$= 1 + 0.035 \times 88.1557 + 88.1557 \times (88.1557 - 60) \times (100 - 88.1557) \times 7.10^{-6}$$

$$= 4.29$$

### 4.3 Determining VoIP capacity of mesh node

Studies conducted to, compare the performance of the leading VoIP protocols in terms of bandwidth consumption and other performance metrics, have shown that VoIP performance is affected by the signalling protocol as well as the codec used [64]. The purpose of the tests described in this section was to determine the capacity of the Alix system when used as a VoIP server. The two major constraints on the Alix systems are memory and processor capacity.

A VoIP call consists of two phases, the signalling phase and the conversation phase during which media flows in real-time takes place. The tests carried out here focus on the conversation phase of the VoIP call. The objective of the test is to determine the number of concurrent *SIP* calls the specified device is capable of successfully handling before the call quality deteriorates.

Requirements:

- Asterisk server on the device under observation
- Call generator to make calls sent to Asterisk
- Monitoring server to monitor the performance of the Asterisk server

### 4.3.1 Experimental methodology

The latest Asterisk server (version 11.7 at the time) was installed on the Alix system described in section 4.1.1, running Debian Linux. The Call generator was implemented using the Dial-Out option in Asterisk installed on another machine. The Call generator sends calls to the SIP server over the wireless interface. The SIP server sends the call out through one of its wired ports (eth0) back to the Call generator. The Call generator answers the call coming in on the wired interface and plays pre-recorded sound to simulate a conversation as shown in figure 4.2(a) and 4.2(b). The tests carried out here focus on the conversation phase of the VoIP call. The Monitoring server was implemented using *Zabbix* <sup>2</sup>.

Processor and memory load was monitored and recorded under conditions of no activity i.e. only the operating system running (*see the Appendix*). Then measurements were taken with Asterisk server running bare-bone with no calls being handled. Then the measurements were taken with an incremental number of calls in 5 minute intervals.

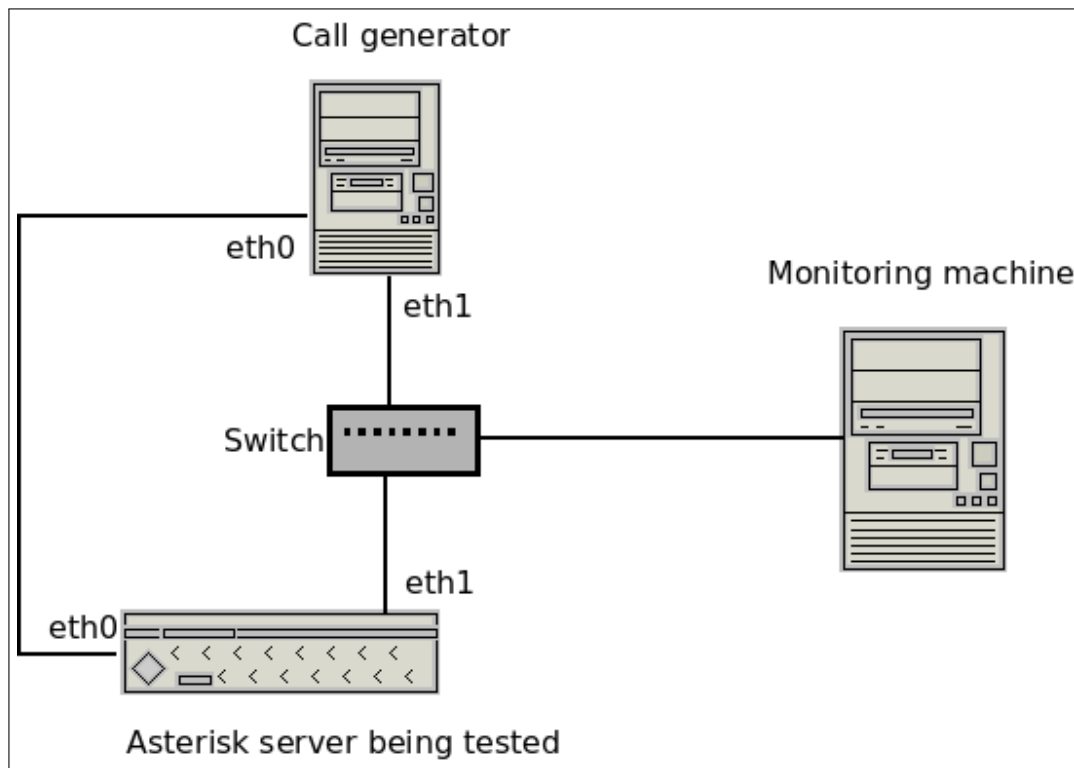
In order to determine the call capacity of the device, the empirical method developed by Modulo Consulting <sup>3</sup> was used. The method is based on a subjective analysis of the audio channel while steadily increasing the number of simultaneous calls.

### 4.3.2 Results of VoIP capacity of mesh node

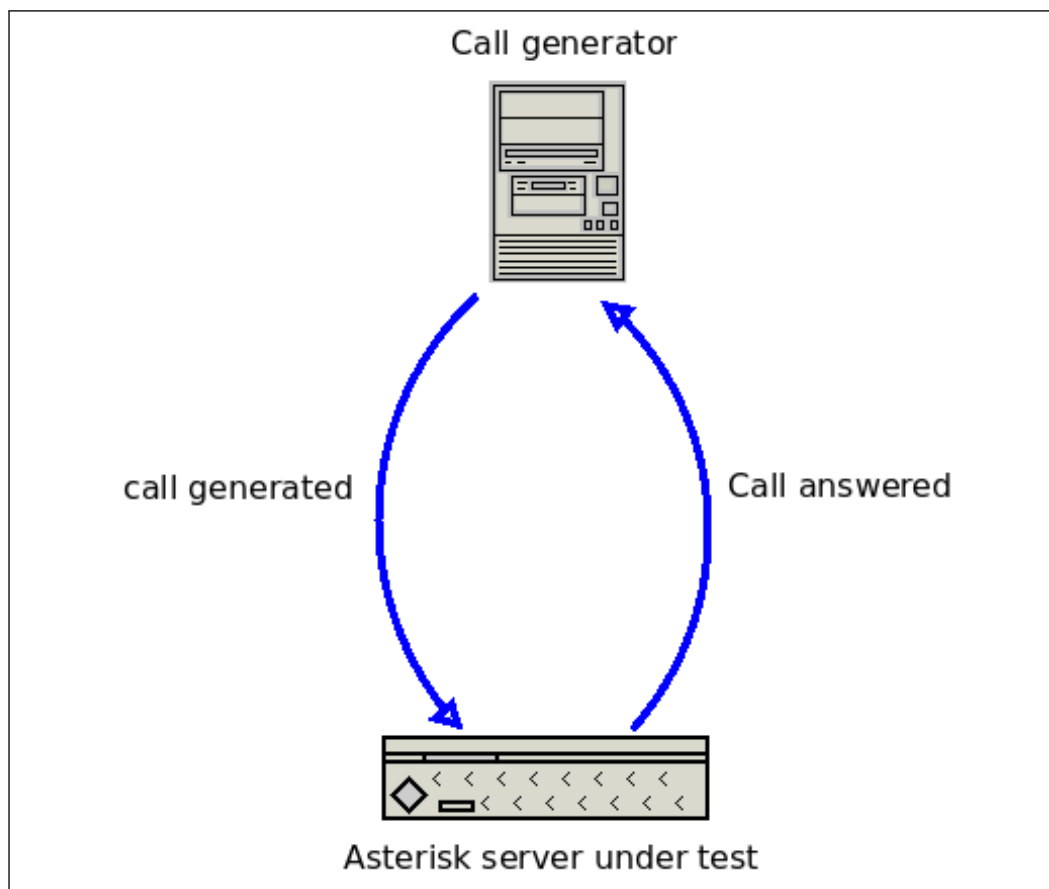
The Alix board when used as VoIP server is capable of supporting up to 15 point-to-point simultaneous calls using SIP and the G.711 based codec. Furthermore, the platform is capable of supporting up to 27 connections in a conference call. These results are based on the Asterisk server running on Alix system which does not have any additional processor and memory load besides handling VoIP calls. Furthermore, VoIP may include multimedia but the tests carried out only considered voice calls. In addition, the VoIP server may sometimes be required to perform protocol and or codec translation for purposes such as relaying a SIP call to a non-SIP callee or integration with PSTN and so forth. The tests carried out involved the same protocol and codec, end-to-end. Therefore, the results represent performance when there is no additional computational and memory load resulting from transcoding or protocol translation processes. Establishing a call involving two different codecs and/or protocols may require a significant amount of computational resources depending on the codec/protocol involved. Therefore, the maximum number of calls the Alix system is capable of handling before the call quality

<sup>2</sup><http://www.zabbix.com/documentation.php>

<sup>3</sup><http://www.modulo.ro/voip-articles-and-tutorials/determine-the-maximum-capacity-of-an-asterisk-pbx>



(a) Physical connections.



(b) Logical setup.

FIGURE 4.2: VoIP stress test setup.



deteriorates is expected to reduce as the node performs other tasks such as routing traffic traversing the mesh network or performing additional VoIP tasks such as protocol or codec translation described above.

## 4.4 Testing the IP-based audio streaming

The IP-based radio station described in the design chapter was implemented on the Alix system. The objective of the tests carried out in this section is two-fold. Firstly, the impact that audio streaming has on the network is investigated. Secondly, objective measurements are carried out to determine the quality of the audio received by the listener.

Two network configurations were considered on the test-bed. The first configuration is the multi-hop single path configuration. This is the type of network configuration where traffic is forwarded over multiple nodes between the streaming server and the listener client. However, there is only one possible path from source to destination as shown in figure 4.3.

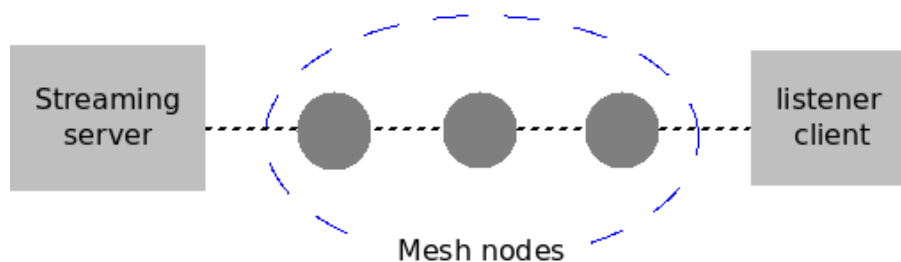


FIGURE 4.3: Audio streaming on multi-hop single-path network configuration.

The experiment was repeated on a multi-hop multi path network configuration. In this case the audio is streamed from server to listener client with multiple possible paths on the network as shown in figure 4.4.

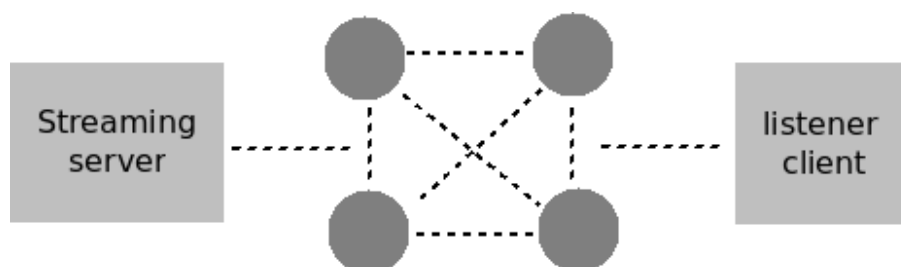


FIGURE 4.4: Audio streaming on multi-hop multi-path network configuration.

#### 4.4.1 Findings

The figure below shows the bandwidth utilisation based on a possible worst case scenario of 1Mbps link capacity. This experiment was conducted under conditions of no other network traffic. It can be seen that the amount of bandwidth available reduces as the number of listener clients increases. The increment is almost linear with each client consuming about 8Kbps. This was experimentally tested on the test-bed with the number of clients up to 20. Based on the observed pattern, it is concluded that the amount of available bandwidth will continue to reduce as the number of clients increase.

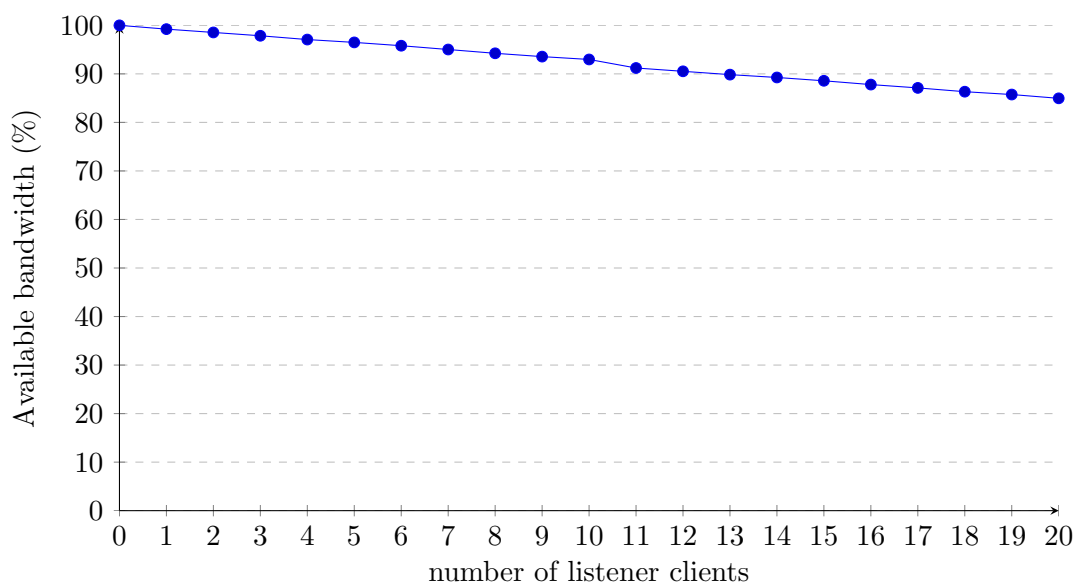


FIGURE 4.5: Audio streaming bandwidth utilisation based on a 1Mbps link capacity.

#### 4.4.2 Evaluation of IP-based radio using PEAQ algorithm

The IP-based radio implementation was evaluated using Perceptual Evaluation of Audio Quality (PEAQ) algorithm. PEAQ is an objective method for measurement of perceived audio quality that is specified in ITU-R recommendation BS.1387-1 [71]. The effects of compressing multimedia for streaming coupled with the network conditions affects the quality of the audio at the listener end. The purpose of this experiment was to determine the impact of network characteristics on the streamed audio quality. Traditionally, testing the audio quality is done through formal listening tests involving expert listener subjects. This subjective assessment approach is not just time consuming but also costly. PEAQ is an objective method of estimating the audio quality as perceived by the user. Subjective listening tests use the Subjective Difference Grade (SDG) as a measure of the Basic Audio Quality (BAQ) expressed as the difference between the signal under test and the reference signal. The output of the PEAQ objective measurement is the

Objective Difference Grade (ODG) which is a measure of BAQ. Table 4.5 shows the ITU-R five grade impairment scale and the corresponding grade and audio quality.

Impairment	SDG	Grade	Quality
imperceptible	0	5	excellent
perceptible but not annoying	-1	4	good
slightly annoying	-2	3	fair
annoying	-3	2	poor
very annoying	-4	1	bad

TABLE 4.5: ITU-R five grade impairment scale in reference to audio quality measurement.

The basic principle behind the PEAQ method is to extract features from the original audio signal before processing for streaming and compare it with the features extracted from the audio signal after processing and streaming across the network in order to estimate the quality of the audio as illustrated in figure 4.6. This is helpful in assessing the impact of effects such as compression techniques, network conditions on the quality of the output audio signal.

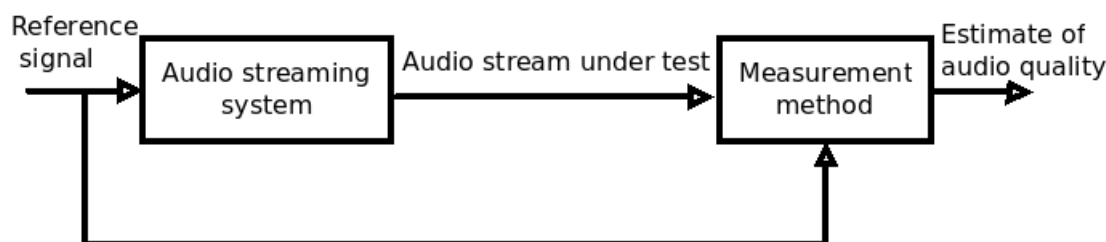


FIGURE 4.6: Basic principle behind objective measurement of audio quality.

The PEAQ method takes in two inputs which are, the unprocessed reference signal at the source and the processed signal under the given conditions that has to be tested. The algorithm determines the distortion in the signal under test and the intermediate output of the signal processing are Model Output Variables (MOV). Further processing of the MOVs, is performed to map the MOVs to the main output of the method which is an ODG, a measure of audio quality ranging from 1 to 5 i.e. bad to excellent. In principle, the PEAQ algorithm measures the audio quality of a signal under test, which has been changed in some way, in relation to a supplied reference signal, which is the unchanged signal.

PEAQ comes in two flavours: a basic version and an advanced version. The basic version is focused on achieving efficient real-time implementations. On the other hand, the advanced version is aimed at attaining the highest achievable accuracy. The additional accuracy inevitably increases the computational complexity of the advanced version in comparison with the basic version.

Before employing a PEAQ algorithm implementation, it is important to verify conformity of the implementation to the ITU-R recommendation. The ITU-R recommendation BS.1387-1 provides a set of *WAV* files with which to test an algorithm implementation. If an implementation reproduces the given Distortion Index (DI) values with tolerance of less than 0.02, it is said to conform to the recommendation. If an implementation does not produce the values with the stated tolerance then it is said to not conform to the recommendation.

For this evaluation, the basic PEAQ algorithm implementation was used. The audio files before encoding at the streaming server were used as reference signals. Because of the inherent difficulty in applying the PEAQ implementation to an audio stream in real time, the audio streaming output had to be captured off of the listener client sound-card and then input into the measurement method. When running the test, the signal under test and the reference signal are time aligned so much so that the effects of coding and network conditions are reflected in the audio being tested.

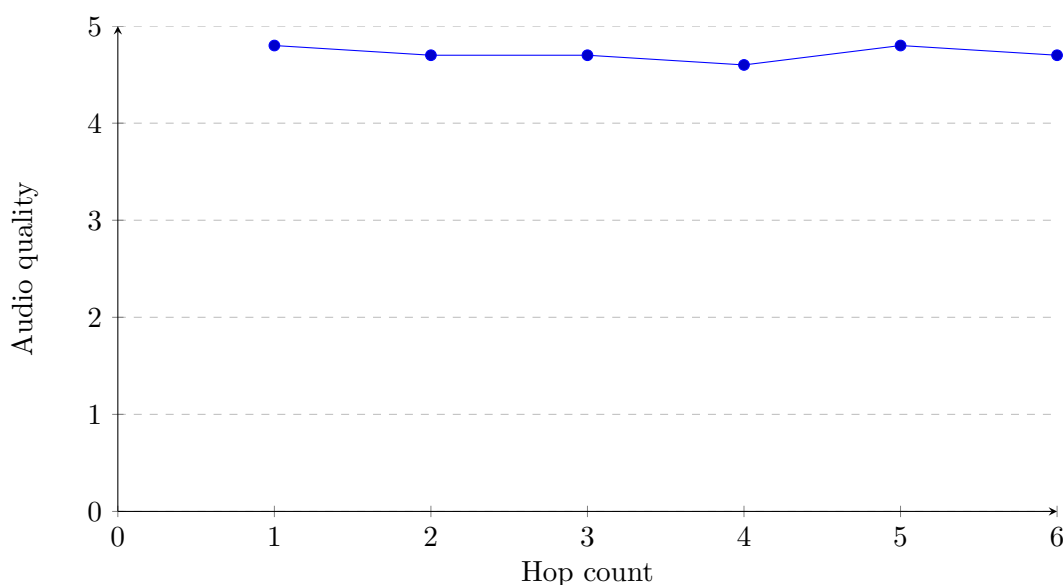


FIGURE 4.7: Investigating the effects of network hop count on audio quality using ogg audio format at 44kHz sampling rate.

For the number of hops investigated, hop count does not seem to have a significant effect on the quality of the audio streamed as observed in figure 4.7. Furthermore, except for the mp3, which shows a consistently poor sound quality at various sampling rates, all the other codecs show a non-linear increment in audio quality as the sampling rate increase with the best results achieved at 32kHz and above as shown in figure 4.8.

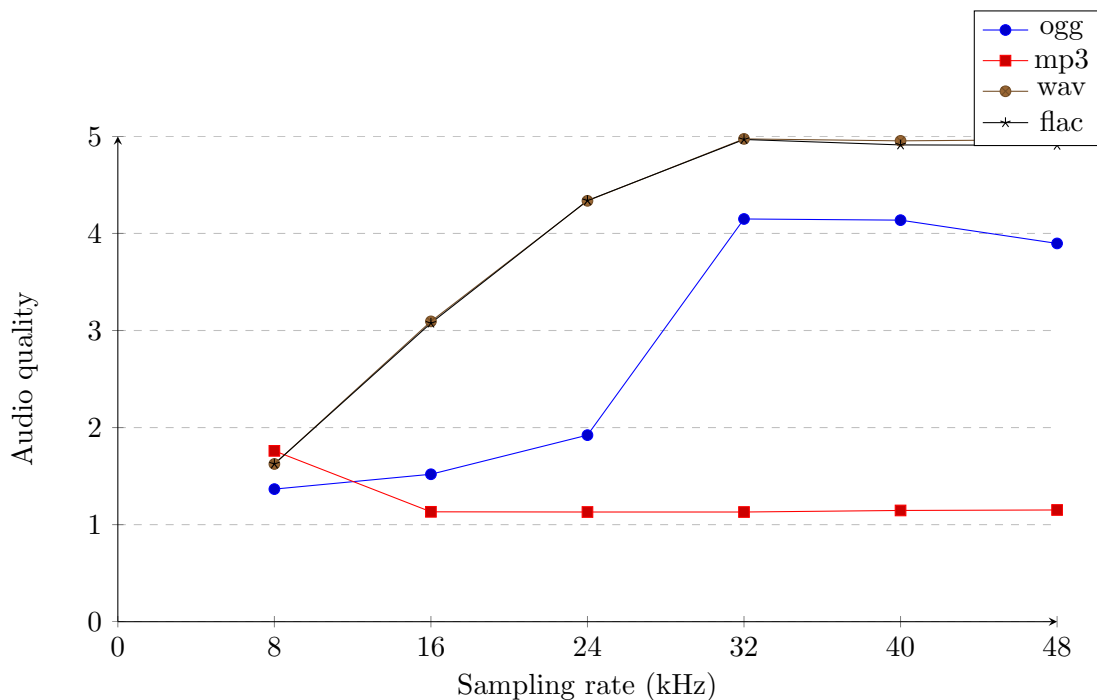


FIGURE 4.8: Investigating the effects of sampling rate on audio quality caused by the common codecs.

## 4.5 Testing the text to voice translation gateway

Files that have to be automatically played by Asterisk are saved as *WAV* encoded in 16-bit, 8000 Hz, mono format. This implies that for every byte of text, approximately 1778.1 bytes of storage capacity is required for the equivalent audio files. Furthermore, a byte of text results in audio file of play duration approximately equal to 0.1 seconds.

There is a 3.5 kilobyte limit on the amount of text that can be synthesised at a time. The heap size may be resized but making it large increases the time required for garbage collection and subsequently increases the time needed to complete the synthesis. This means that for large amounts of text the speech synthesis has to be performed in parts. The resulting audio files can then be concatenated or played back consecutively.

Furthermore, the translation gateway mechanism relies on the maintenance of information in audio format equivalent to what is in textual form. Text files are subject to change from time to time. In order to keep the audio files up to date, changes in text files need to be tracked and the associated audio files updated accordingly. There are several options available in monitoring the files to detect whenever there is a change. The challenge is to monitor and detect the changes in the most efficient manner.

The first option is to keep a list of text file names along with the last-modified date. The system can then poll the file system and compare the last modified date with

what is on record, and if there are discrepancies, re-synthesise the text to speech and overwrite the associated audio file. However, this method proved unreliable because of the way functions return the last-modified time. Sometimes even when a file has not been modified, the system function returns a value different from what was record. This may be explained by the fact that the value is a float with inevitable rounding required causing inconsistencies in the values. In addition, manually polling the file system suffices for a small-sized file system but degrades performance when the number of files increases.

The other alternative is to let the kernel handle the monitoring and have the notifications exported to user space through system calls. This can be achieved using a module such as *inotify*<sup>4</sup>. The implementation for this dissertation used *pyinotify*<sup>5</sup>, which is a Python version of the module. The objective is to monitor the file system and trigger notifications whenever there are changes for purposes of ensuring that the system maintains updated audio files of information. Several events to monitor can be specified using *pinotify*. The most relevant event for this work is the *IN-MODIFY* event. This event triggers a notification whenever a file is modified. One shortcoming of triggering the text-to-speech process based on this event is that, when monitoring a directory, it is very difficult to return the name of the file causing the event. The alternative is to use the *IN-CLOSE-WRITE* event. This event is triggered whenever a file is written to. Using *IN-CLOSE-WRITE* the name of the file causing the event can easily be returned. This is important because the text-to-speech processing mechanism uses this file name as target for the source of text to synthesise to speech. However, from the tests carried out, the behaviour seems to depend in part on the text editor used in writing to the file. For instance, using the 'echo' command, the trigger behaves as expected whereas when the file is edited using *gedit* for example, the event is triggered twice for every one time.

#### 4.5.1 Evaluation of the synthesised speech using PESQ

This section describes the experiments which were carried out to asses the quality of speech synthesised from text. Synthetic speech is the speech produced by the text-to-speech system. The evaluation was done by measuring the speech quality using Perceptual Speech Quality (PESQ). PESQ is an objective method of assessing speech quality for narrow-band telephone networks and speech codecs described in ITU-T recommendation P.862 [72].

---

<sup>4</sup> An API that provides a mechanism for monitoring file system events: <http://man7.org/linux/man-pages/man7/inotify.7.html>

<sup>5</sup><http://pyinotify.sourceforge.net/>

The PESQ method aims to predict the perceived quality that would be given to the signal under test by participants in a subjective listening test. The method is applicable to speech codecs as well as end-to-end measurements. The ITU-T recommendation P.862 describes the PESQ method in which an original signal  $X$  is compared with a degraded signal  $Y$  that is the result of passing  $X$  through the communication system. The process involves transforming the two input signals to the PESQ method into an internal representation that models the human auditory system. Using this principle, the quality of the speech automatically generated from text by the system was tested using human recorded speech as the reference signal. The PESQ output is then mapped onto the subjective MOS scale in order to objectively estimate the speech quality that would be perceived by users of the voice-based system.

PESQ was developed primarily for use in telecommunication systems however, experiments conducted by researchers [73] show that there is a high correlation between the objective PESQ value and the MOS scores obtained from subjective tests of synthetic speech quality assessment. It is on this basis that the PESQ method was deemed appropriate for this dissertation. In addition, the audio signal under evaluation was obtained from across the network so as to ensure that the impairments introduced by the speech synthesiser as well as the conditions of the communication network on the speech quality are captured.

The Mean Opinion Score (MOS) scale (described in ITU-T recommendation P.800) is usually used in subjective methods of speech evaluation. During the test, the listener grades the speech under evaluation and assigns it to one of the categories shown in table 4.6.

MOS value	Meaning
5	excellent
4	good
3	fair
2	poor
1	bad

TABLE 4.6: MOS scale categories

But, as pointed out earlier, PESQ is an objective method based on perceptual modelling. The *time alignment* routine in the PESQ algorithm provides the perceptual model with delay values to allow corresponding parts of the reference signal and the signal under evaluation to be compared. The audible difference between the reference signal and the signal being tested is calculated from the internal representations of the two signals. The output of the PESQ algorithm implementation is expressed in terms of *MOS-LQO* (Mean Opinion Score - Listening Quality Objective). MOS-LQO specified in ITU-T

recommendation P.862.1, allows the raw PESQ scores to be linearly comparable with the MOS scale. MOS-LQO ranges from 4.5 to 1.0 where 4.5 is the best and 1.0 is the worst.

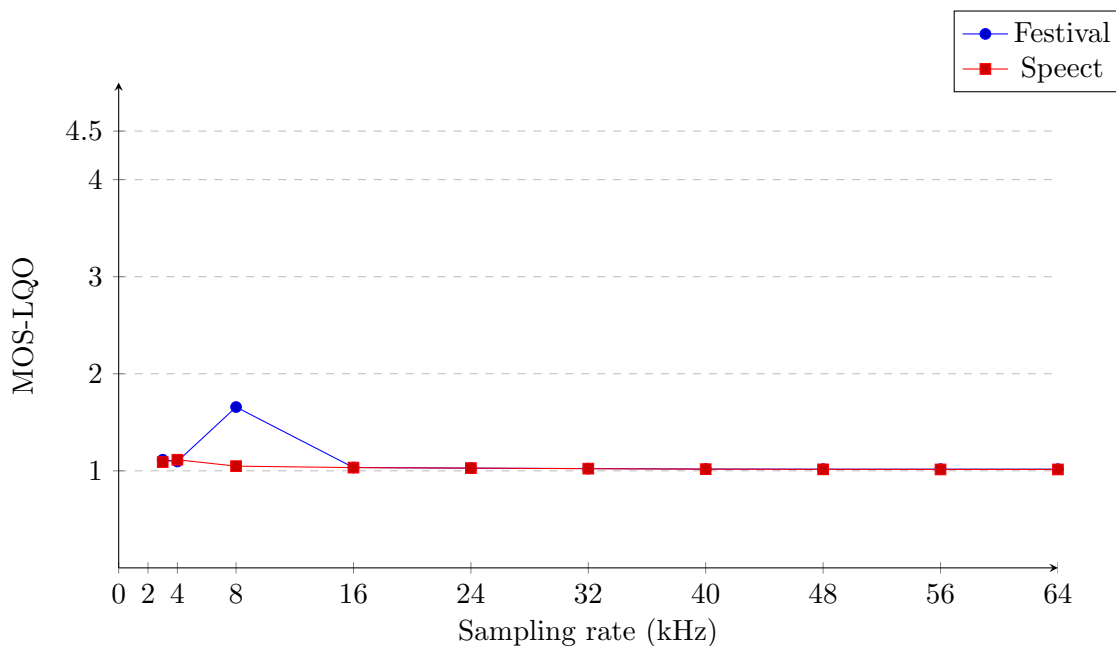


FIGURE 4.9: Assessing the quality of automatically generated speech.

Some conclusions can be drawn from figure 4.9, which shows the obtained MOS-LQO scores for two available TTS engines, *Festival* and *Speect* used in the automatic mapping of text to speech. Firstly, the effect of up-scaling the output audio from the TTS was investigated. From the experimental results obtained, up-scaling does not enhance the quality of the automatically generated speech as can be seen in figure 4.9. Increasing the sampling rate from 8kHz onwards shows no significant difference in the resultant MOS-LQO score. Furthermore, based on the PESQ method output, Festival exhibits better performance at 8kHz but there is no noticeable performance difference from that point onwards between the two TTS engines.

#### 4.5.2 Some limitations of the translation gateway

By experimenting with the translation gateway using the speect text-to-speech engine, the following limitations were identified:

- Synthetic speech sounds different from natural speech and some users are bound to find it less appealing.



- The system lacks extensive Natural Language Processing (NLP) capabilities and as such, the text to be synthesised has to be normalized i.e. plain text. For instance, '*1 day*' needs to be '*one day*'.
- It is difficult to map pictures appearing in text into equivalent voice form. However, this can be circumvented by embedding sufficiently descriptive *Alternative text* for images in the markup tags and then providing an audio description of the picture based on the Alternative text.

## 4.6 Summary and implications of the results

A MOS score above 4 i.e. high quality of mobile VoIP has been achieved. However, the maximum call capacity of 15 simultaneous calls imposed by the Alix system board implies that the Alix system cannot be used as a VoIP server where a larger number of simultaneous calls needs to be serviced. In this case, a decentralised model would be ideal. If a centralised model is desired, then a high performance computer is required for the VoIP server.

Regarding the translation gateway, because of the inherent challenges with the current state of text-to-speech technology for static information, human recorded transcription of the information is needed for an enhanced user experience. The automatic text-to-speech system can be used supplementary for dynamic textual information. Furthermore, due to the other inherent limitations of text-to-speech systems such as inability to identify paragraphs, it is best to include a *pre-synthesis phase* in the translation process. The text-to-speech engine takes all the text passed to it and converts it to speech. The purpose of this pre-synthesis phase is to analyse and normalise the text before passing it to the text-to-speech engine. Among the tasks performed during this phase include, selecting the text to synthesise which involves identifying and eliminating other portions of text which might be used for formatting purposes as is the case with an HTML file. By including this pre-synthesis stage, paragraphs can be identified and appropriate pauses and silence included in the synthesised speech thereby enhancing the quality of the resulting audio format of the information.

## Chapter 5

# Discussion on the findings

In view of the potential benefits proffered by enhanced communication and access to information, this dissertation set out to determine the technical feasibility of implementing such services using the available infrastructure under the constraints of a rural context. In this section, the research questions posed in section 1.2 that have been answered through the outcome of this dissertation are revisited and a discussion on some of the findings is included.

### 5.1 How can the WMN be used for mobile VoIP?

The aim was to enable users make VoIP calls using mobile phones across the WMN infrastructure. This research question was of paramount importance because, by adding support for mobile VoIP we are improving communication. In addition, the prevalence of mobile phones implies that adding support for these devices will open the door for other voice-based phone services. The high cost of and sometimes non-existent cellular network service in rural areas is the major driving force behind the search for alternative communication means.

The results of the test-bed experimentation shows that mobile VoIP support can be achieved using a centralized architecture. Using the E-model, we were able to achieve a Transmission Rating Factor above 88 (*see section 4.2*), which is deterministically equivalent to a MOS value above 4. ITU-T emphasises that the output of the E-model is meant for planning purposes however, it can be transformed to give estimates of user opinion [4]. Therefore, we are confident that the measure provides a satisfactory estimate of the VoIP call quality.

---

Asterisk meets the requirements of a VoIP server. The centralised architecture means that users have to be defined at the server level. Mobile devices with network connectivity are able to make VoIP calls but, for a call to be successful both the caller and the callee must be registered with the VoIP server. Establishing network connectivity for all users over a large area requires the use of multiple Access Points and the range is affected by among other things, obstacles, presence of interferers, antenna design and radio power settings.

Using a centralised approach potentially creates a bottleneck and a single point of failure at the VoIP server when there is a high call volume . A VoIP server may be said to be overloaded when there is insufficient resources namely, CPU and memory capacity to process incoming connections. An increase in the number of connections may lead to delays in connection establishment and a subsequent increase in the number of unsuccessful connections once the server gets overwhelmed. Overload control mechanisms are required in order to maintain quality of service.

This dissertation proposes a distributed-centralisation approach illustrated in figure 5.1 where client registration and signalling are centralised but, multiple servers are employed for purposes of load-balancing and fault-tolerance. When using Asterisk for the VoIP server, it can be distributed among multiple servers. This server decomposition is made possible by the Inter-Asterisk Exchange (IAX) protocol that keeps the multiple servers functioning as a single atomic unit [3]. Using multiple VoIP servers leads to three possible call scenarios:

- *Intra VoIP server call.* This is when the caller and the callee are registered under the same VoIP server. When a caller dials a number, the VoIP server checks through its registered clients. If the dialled number is a registered client, the call is established between the caller and the callee.
- *Inter VoIP server call.* This scenario arises when the caller and callee are registered with different VoIP servers. When a caller dials a number, server 1 interacts with server 2 to verify callee registration. If the callee is registered with server 2, the call is established between the two clients.
- *Inter community call.* This case arises when a client from the community needs to call a number outside the VoIP network. In this case, connection is established via a gateway to PSTN (See section 2.3.2 for PSTN gateway implementation).

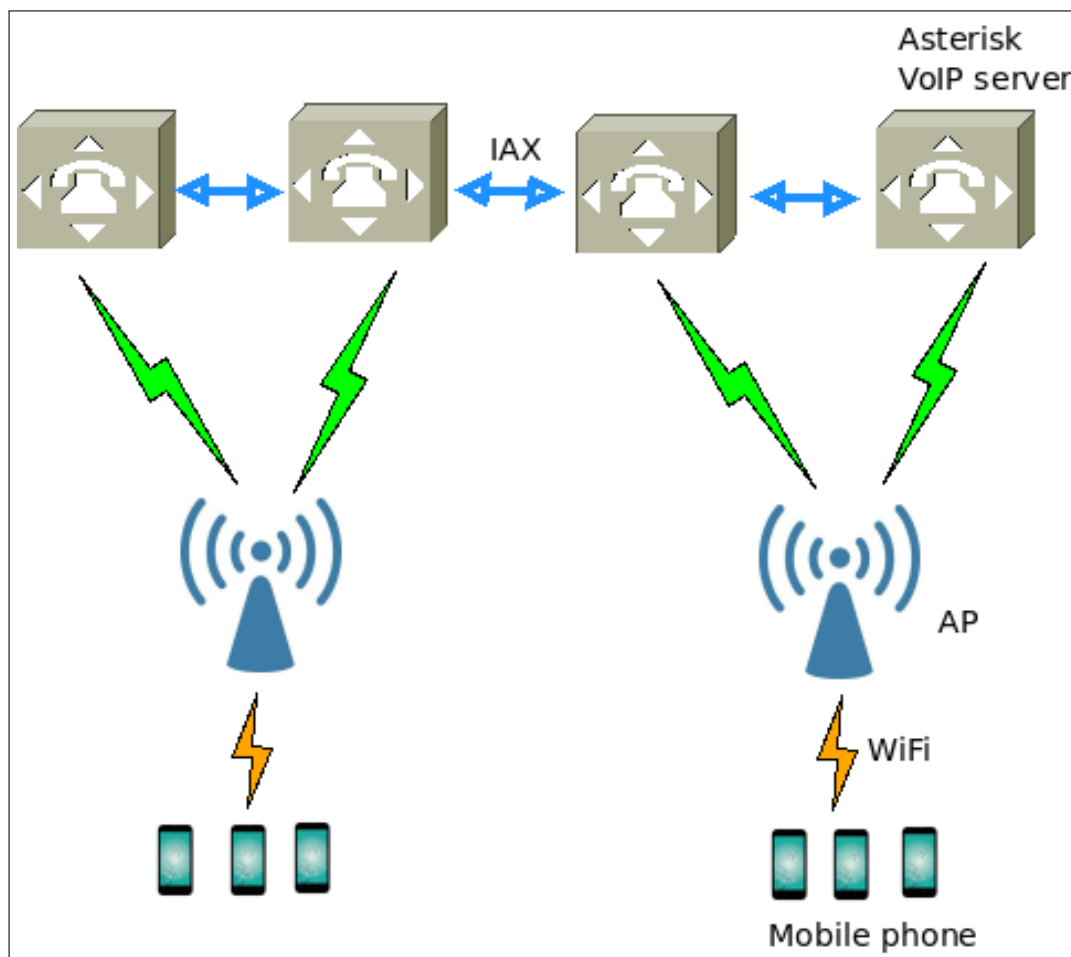


FIGURE 5.1: Distributed-centralised mobile VoIP architecture.

### 5.1.1 Factors affecting the quality and capacity of mobile VoIP

When a centralised architecture is used, the VoIP capacity might be determined by the VoIP server's processing and memory capability. In section 4.3, a test was conducted to determine the VoIP server's capacity to handle multiple simultaneous calls. This was important to find out because, once the maximum successful number of simultaneous calls is established, mechanisms can easily be built into the server to drop additional connections and appropriately notify users trying to make calls after this limit has been reached. The principle idea is to avoid exceeding the server's capacity because any attempt to do so is likely to degrade the VoIP quality of all initially successfully established connections. To explain further, suppose the server is capable of successfully handling a maximum of 15 simultaneous calls and CPU capacity is the limiting factor, an attempt to add the 16<sup>th</sup> call overwhelms the server thereby affecting the quality of all the other 15 previously successful connections.

In addition to the VoIP server requirements, interference and the lossy nature of the wireless transmission medium tends to affect the end-to-end quality of the wireless based

---

VoIP service [74]. Mesh nodes in a multi-hop environment placed at different places of the coverage area, suffer interference of varying degree at different times.

Furthermore, consider an 802.11(a/b) network cell supporting data rates of up to 11 Mb/s and a full duplex VoIP connection requiring, 128 Kb/s for instance, one would expect simultaneous calls in the order of 85 to be supported. In reality, the effective capacity is far less because the MAC layer implementation places additional capacity constraints. Clients connected to the same AP use the same channel to transmit but, for successful transmission only one is allowed at a time. The MAC protocol is intended to prevent collisions from happening when multiple clients using a shared medium attempt to transmit at the same time. There are two techniques used at the MAC layer. The first one is called Point Coordination Function (PCF) and the second one is called the Distributed Coordination Function (DCF). PCF is where the AP makes decisions on which of the connected clients gets to transmit at a given time. On the other hand, DCF belongs to the CSMA/CA suit of protocols used in wireless networks where channel access is determined by the clients with no central coordination mechanism in place. The nature of DCF makes it more susceptible to media access DoS attacks. While PCF offers an alternative for reduced DoS attack susceptibility, not all IEEE 802.11 hardware supports PCF. There is also the Hybrid Coordination Function (HCF) which is a federation of DCF and PCF. The MAC mechanism in place to ensure that collisions do not occur when clients transmit using a shared medium leads to increased delays as the number of clients increases. This is because clients contending for the transmission medium have to wait much longer for a turn to transmit. This hunch is confirmed by several studies [74], [75], [76]. The delay at the MAC layer may not be a big issue for delay tolerant applications such as data transmission. However, for VoIP, an increase in packet delay results in significant degradation in the quality of service because of the delay intolerant nature of the service and the stringent real-time requirements.

Added delays on the communication channel may be mistaken for turn-taking pauses, which leads to loss of synchronicity in the conversation and potentially alter the perceived meaning of a message [77]. For instance, if one party asks a question, the delayed response caused by processes on the communication channel may be construed by the other party as hesitation. The ultimate meaning of the message becomes subject to people's interpretation of hesitation such as it being a sign (or a lack) of openness, confidence, or honesty.

This implies that with the centralised VoIP approach, when the VoIP connections originate from different network cells, the VoIP server's processing capability limits the total number of simultaneous calls. However, when the connections originate from the same 802.11 cell, the MAC implementation places added restrictions on the maximum number

---

of clients connected to an AP that can take part in VoIP instances at the same time for a given level of QoS. In addition, the transmission rate reduces as the client moves farther away from the AP. Therefore, the VoIP capacity and quality are further affected by the distance of the clients from the AP.

One possible quick way to increase AP capacity is to add more APs in the area. However, for multiple APs in range of each other to not interfere with each other, they need to operate in non-overlapping channels. The requirement for APs to operate in non-overlapping channels for co-existence restricts the amount of capacity increase achievable simply by increasing the number of APs because of the limitation on the total number of IEEE 802.11 usable non-overlapping channels [76].

## 5.2 How can the WMN infrastructure be used to disseminate community information?

From the literature review conducted in chapter 2, improved access to information as well as information sharing among the target users were identified as areas requiring enhancement. Based on the numerous reports (*see section 2.4.1*) on the benefits of employing sensor technology in agriculture, it was thought that providing information about the condition of the environment would benefit the predominantly small-scale farming communities. In addition, the mobile phone was identified as the most prevalent technology among the target population. In response to the above needs, the following services were identified:

- **Environmental monitoring**

WSN allows sensors to be spread over the field thereby enabling distributed measurements. The information collected about the environmental characteristics of the area of interest provides the basis for decision making viz. irrigation decision (when and how much water to apply), pesticide application, etc. The information obtained in real-time enables early warning and allows the farmer to take a timely control action. The deployment of sensors also provides data in electronic format which might be useful for agriculture experts. With the monitored parameters, Agarkar et al. [78] places the possible analyses into categories as follows:

- *Abiotic stress analysis*
- *Micro-climatic analysis*
- *Crop nutrient analysis*
- *Crop disease analysis*

---

As an example of the relevance of these analyses, consider crop nutrient analysis: plants require specific nutrients in sufficient amounts for their growth. Most of the soils lack the nutrients in sufficient quantities hence the need for supplements through various methods such as fertilizer application to the soil. Given suitable ground sensors, WSN can help in characterising the nutrient status of the soil and aid the decision regarding when, how much and what kind of supplement to add. Furthermore, models such as the one developed by Broom et al. [79] have been developed that can be used to analyse the data and determine the *infection index* for crops such as grapes. The infection index is used to determine when a crop is at risk. With crop disease analysis, instead of applying pesticides based on some fixed schedule, farmers are able to make the application when and where required.

- **Enabling phone access to online textual information**

A lack of Internet connectivity coupled with limited textual literacy limits people's access to information in rural areas. The idea was to take advantage of the mobile phone penetration but, depart from the one-to-one voice communication usage model.

Answering this question needed a contextualised approach in the design of the system. In response to limited textual literacy and a lack of Internet connectivity, the idea was to develop a voice-based information system. With the help of an IVR (described in section 3.2.3.3), users are able to easily interact with the system via a voice-based interface. In addition, by using a translation gateway (described in section 3.2.3.4) that is capable of mapping a given phone number to a URL and capable of converting textual information to a voice format, mobile phones can be used to access online information. Using the phone for online information access ensures access for a larger population because of the extent of penetration of the technology. Using a voice interface has many advantages such as being implicitly learnt because of the substantiality of spoken communication in everyday life. However, voice based interfaces have a number of usability issues such as the linear nature of the interface, which tends to be time consuming in comparison with visual interfaces where options can be presented in parallel. In addition, voice interfaces require the user to pay attention while using the system, the potential of some users failing to cope with the additional cognitive load implies that generally, the voice-based interface can only hold a limited amount of information.

The PESQ method (described in section 4.5.1) was used to assess the quality of the audio produced by the text-to-speech system. MOS-LQO values barely above 1 (i.e. bad) were obtained. There are two possible explanations to the poor PESQ scores of the translation gateway. Firstly, the apparent poor performance may be caused by a lack of robustness in the time alignment module of the PESQ

implementation as the comparison is made between the reference signal and the synthetic speech. Secondly, when evaluating TTS using PESQ, ideally the voice used in the reference signal should be the same one that was used in the TTS engine implementation but, the experiment used a reference signal which is not the same as the voice used in the TTS engine. This immediate natural difference in the voices between the two input signals into the PESQ method may account for the poor MOS-LQO values obtained.

Therefore, it can be seen that the PESQ method of evaluation may not eliminate the need for subjective tests but, it provides a way of automating the testing process thereby allowing the testing of the different system configurations and obtain the results quickly.

- **Enabling one-to-many voice communication**

By developing a phone-accessible voice-based forum (described in section 3.2.3.1) which allows users to post voice messages and comment on others' posts, farmers are able to share knowledge with other farmers. The forum approach where users themselves create their own content ensures that the content is kept locally relevant which is a key concern in ensuring that the system gains widespread adoption.

- **IP-based radio**

Using PEAQ objective method (described in section 4.4.2) of evaluation, streamed audio quality score above 4 i.e. very good was achieved. Increasing the sampling rate increases the audio quality but this is at the expense of increased bandwidth utilisation due to a subsequent significant increase in the file size. The ogg file format leverages audio quality against file size. From the experiments carried out, increasing the sampling rate or up-scaling beyond 32kHz shows no significantly noticeable improvement in the audio quality despite a continued increase in the resulting audio file size.

Bearing in mind the technical challenges faced by rural communities, the task was to implement the above services and then answer the question:

### 5.2.1 Can the services be implemented solely on low-power devices?

Service quality is often times expressed in terms of the mean opinion score (MOS). A MOS score above 4 was achieved for the proposed services, except for the automatic translation of text to voice that had a score of 1.6, i.e. poor. But, this poor performance cannot be attributed to the hardware properties of devices as discussed in section 5.2, second part. Based on the findings of the measurements using objective methods, the conclusion is that the services can successfully be implemented using low-power devices.



---

### 5.3 What is the impact on overall network performance of adding the new services on the WMN?

#### Node level performance analysis

The WMN comprises mesh nodes that forward traffic from source until it gets to the destination. Investigating the impact on network performance is an important undertaking because, the implementation of an IP-radio on the community based network was considered as a tool for information dissemination. Multimedia transmission is generally known to be resource intensive.

Streaming audio in real-time requires high transmission capability in order to maintain service quality. While transmission capacity is not a huge concern in static wired networks, it is a major concern in wireless networks and much more so in a multi-hop environment. The objective of the experiments carried out was to firstly determine the quality of service of the IP-based radio. This was done by analysing important streaming metrics namely, throughput, packet loss and jitter in a multi-hop network topology. Secondly, given the current network configuration (single radio, single channel) and the service architecture, we set out to assess the impact the audio streaming would have on the quality of VoIP service in terms of bandwidth requirements the services have. More specifically, suppose one client at the end of the network connects to the audio streaming server and another one decides to place a VoIP call, will the call be successful? In addition, how does the case change with an increase in the number of clients streaming audio and other VoIP calls being made simultaneously? Experiments were conducted with multi-hop single-path and multi-hop multi-path network configuration. The resource consumption at the nodes as the audio is continuously streamed across the network was investigated.

#### Network level performance analysis

Continuing from the previous research question, the WMN consists of nodes connected in such a way that each node is linked to at least two other nodes. This ensures multiple possible paths from source to destination. The nodes route traffic through the best link as determined by the routing protocol. The path redundancy characterising WMN ensures that nodes are able to route traffic via alternative links when any one of the nodes go down. However, once a node runs out of memory for example, it will begin to drop packets thereby potentially affecting network performance. The aim was to determine how the overall network performance might be affected by the introduction of the service. Network parameters that might significantly affect the performance of the VoIP service are latency, jitter, packet loss, and data rate.

---

Latency is the amount of time it takes a packet to move from source to destination on the network. An extreme increase in the transmission time will make it difficult to conduct an interactive conversation. Jitter describes the variation of latency on a given set of packets. Though jitter may go unnoticed in data networks because of TCP/IP handling, it can have a very significant effect in a VoIP environment. Jitter can be minimised by employing buffers at the destination for rearranging packets that arrive out of order. Packet loss is the ratio of packets received to packets sent. Packets may be lost in transit due to congestion or may be discarded at the destination if the TTL expires. Packet loss will result in loss of speech frames.

Simply put, a network is said to have good performance if the latency, jitter and packet loss are low. For a maximum of 16 mobile devices available for use on the test-bed, we observe no decline in the obtained MOS scores for the quality of service. A combined average of transport, propagation and jitter buffer delay of 1.2355 ms and average packet loss of 1% were observed. The performance was deemed satisfactory based on the achieved metrics, which are well within ITU recommended range.

By and large, the experiments were carried out in a controlled indoor environment. We would generally expect a difference in the results obtained from outdoor WMN deployments. This is because the performance of individual nodes is influenced by the location in terms of the electromagnetic interference received. These variations over time tend to fluctuate the quality of the wireless links between the source and the destination.

### 5.3.1 Recommendations

Tests of the IP-based radio implementation (shown in section 4.4.1), show an increase in bandwidth consumption proportional to the number of connected listener clients. This can be explained by the fact that *unicast* had been used. Unicast is where packets are sent to a single recipient specified by the destination address. When a client connects to the audio streaming server, the server transmits packets to the client. When another client is connected, a separate communication relationship is established and the server begins to send packets to the client. This implies that for  $N$  clients, the server has to pack and send the same packets  $N$  times i.e. separately for each of the connected clients as illustrated in figure 5.2. For a media streaming service like IP-based radio, unicast results in multiple transmissions of the same content and subsequent inefficient utilization of network bandwidth.

When the number of listener clients increases, IP-based radio using unicast may bog down the network, causing increased traffic delay and possibly rendering the network unusable for anything else. Therefore, for audio streaming targeting a large number of

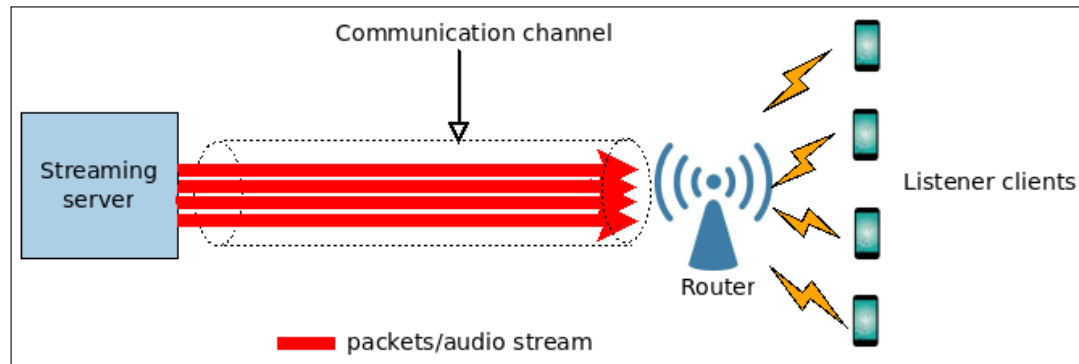


FIGURE 5.2: Unicast audio streaming bandwidth utilisation.

users, it is advisable to adopt a *multicast* approach. Multicast makes efficient use of resources in one-to-many transmissions by eliminating redundant packet transmission. Instead of sending packets to each of the clients, the streaming sever sends packets to a multicast address and the transmission is received by all the clients participating in that particular multicast group. A multicast address is any class D address i.e. an address in the range 224.0.0.0 to 239.255.255.255. Thus, using the multicast approach, the streaming server established only one communication relationship with the multicast group encompassing all the member clients as illustrated in figure 5.3. Multicast makes efficient utilisation of server and network resources in that for  $N$  client connections, the server packs and sends the packets only once.

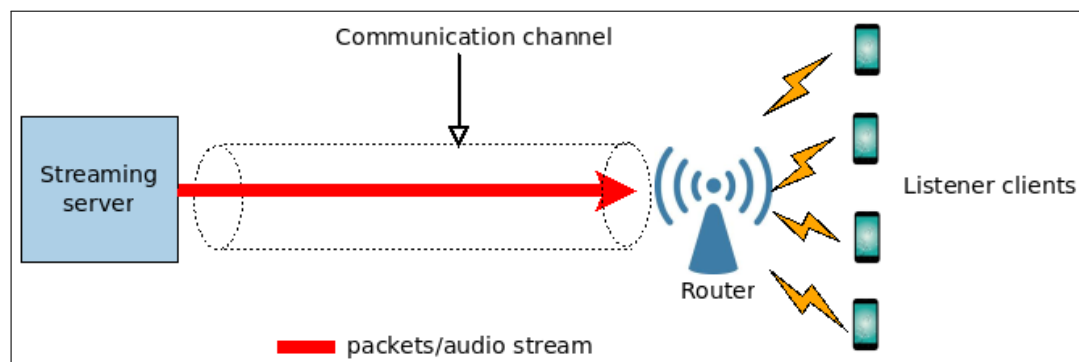


FIGURE 5.3: Illustrating bandwidth utilisation for audio streaming using multicast.

Using multicast to stream multimedia targeting a large number of recipients offers many advantages such as reduced CPU load on the streaming server and efficient bandwidth utilisation. However, there are some drawbacks to the multicast approach namely, it becomes impossible to see who is connected to the server at a given time. Furthermore, the server continuously streams to the multicast address thereby consuming network resources even when there is completely no client connected. Multicast can be implemented at the *data link layer* (where packet replication takes place at the Ethernet switch), *network layer* (where replication takes place at the router) or *application layer* (where replication is handled by the end hosts).

## Chapter 6

# Conclusion

This dissertation looked at the enhancement of information dissemination, and communication in rural areas. Chapter 3 presented the system framework designed for communication and improved access to information, taking into account the challenges of system implementation in a rural community. The distributed nature of wireless mesh networks supports many applications and services besides Internet access dominant in urban communities. This research considered mobile VoIP, IP-based radio station, and translation gateway for text-to-voice conversion as examples of such services aimed at improving communication and the dissemination of information in the community. Based on the achieved quality of service assessed using objective measurement methods, it is concluded that these services can be implemented using low-power devices. However, for the IP-based radio, further optimisation techniques need to be researched in order to stream audio on a continuous basis using low-power devices due to main memory constraints. In addition, more work is needed in the implementation of the translation gateway to enhance the quality of speech automatically generated from text.

Furthermore, mobile phones have primarily been used for one-to-one voice communication. Lack of education coupled with textual illiteracy limits access to information for the majority of people in rural areas. This dissertation took these constraints into account in the design and implementation of the information system. The use of the mobile phone to access online information and to conduct one-to-many asynchronous voice communication (described in detail in chapter 3 section 3.2.3.1) has been considered. In order to serve textually low-literate users, there is optimism that the voice-based system holds the answer.

In addition, service providers may not generate as much revenue from service provision in rural communities as they would from urban areas because of the low payment capacity of users that results from low household incomes among the rural population. Therefore,

low-cost communication solutions are to be considered as a development tool rather than a commercial venture. An ideal system serving a rural small-scale farming community ought to be affordable, and self-sustained.

There has been efforts from the research community aimed at making the deployed solutions viable and financially self-sustained by applying business models that are specially designed for these communities. For instance, Dabba<sup>1</sup> allows its members to make free VoIP calls to members within the network, but charges for calls made to the outside i.e. to PSTN or mobile phones. The revenue generated from the charged calls is used to sustain the system in terms of maintenance and expansion costs. However, Roro .et al [80] argue that such a business model would not work in communities where nearly all the calls are intra-community. They propose a participatory approach in which the community is let to define the rates and method of collecting money from the beneficiaries of the system. The money collected is then used for purposes such as maintenance and upgrades.

By and large, there may not be a universally ideal business model for rural community based projects, each possible model has its set of pros and cons. Needless to say, it is important to take time to understand the community and determine a business model that best suits the community. Furthermore, it has been revealed that the existence of a technology does not imply a successful deployment, the community's internal politics and other factors tend to come into play at various stages [36].

In conclusion, the circumstances surrounding rural communities need not take away the potential benefits that ICT holds. By using low-power devices and designing user interfaces that suit the culture and cognitive abilities of the rural population, ICT services can be delivered to these under-served areas thereby bridging the digital divide. This dissertation focused on establishing the technical feasibility of delivering the proposed services facilitated by low-power devices. The possible long term social and economic impact of introducing such services might not be estimated at this stage, however, the technical aspects of implementation have been investigated and satisfactorily tested. The author looks forward to the adoption of the proposed wireless mesh network system services in rural communities.

---

<sup>1</sup>Dabba Telecommunications, a South African voice and Internet company that claims to specialise in delivering service to multi-tenant dwellings and township neighbourhoods. <http://dabba.co.za/dabba/Home.html>

## 6.1 Future work

The dissertation focused on the technical requirements to ascertain implementation. Immersion into the target community is required in order to identify specific information needs. In addition, more work is needed in exploring techniques for making voice-based interfaces more usable, user tests are required in order to uncover culture-specific system usability issues. For instance, given a voice menu, it is unclear in the Human Computer Interaction field whether a long menu list is better than a short nested menu, the ultimate user experience seems to depend on the user's level of education and cognitive ability among other things.

In addition, the current mobile VoIP implementation is limited to WiFi capable mobile phones. Future works would include integrating a GSM air interface in order to extend the service to non-WiFi capable commodity phones.

Furthermore, future work would include a response to the shortcomings of WiFi upon which the current network implementation depends. The departure from analogue to digital television has left some frequencies unused. The White Space Coalition, IEEE 802.22 Working Group and other proposals have advocated the use of white spaces provided by the switch-over to digital television to provide wireless broadband Internet access. Using TV white spaces offers many advantages over the current WiFi technology.

Lastly, further future works would include partnering with experts in fields such as economics and attempt to determine the economic impact of the services being offered by the WMN in terms such as, agricultural productivity and household income.

# Appendix A

## Appendix

### A.1 Sample measurements

This section presents some of the data obtained from the experiments. CPU and memory utilisation by the nodes was observed while none of the services were running. By making this initial observation of node resource utilisation status, when additional processes were initialised on the nodes, CPU and memory utilization attributed to the added processes was easily deduced.

In figure [A.1](#), the CPU is observed to be the bottle neck on the capacity of simultaneous active channels. This is confirmed by the evidence in figure [A.2](#) which shows that there is still a significant amount of memory still available after the node becomes unresponsive to additional calls. Asterisk is characterised by a low memory footprint. The memory consumption for the total number of calls tested is around 35MB. The limited number of calls tested with, makes it hard to determine the per-call memory requirements because of the low memory consumption increment caused by individual calls.

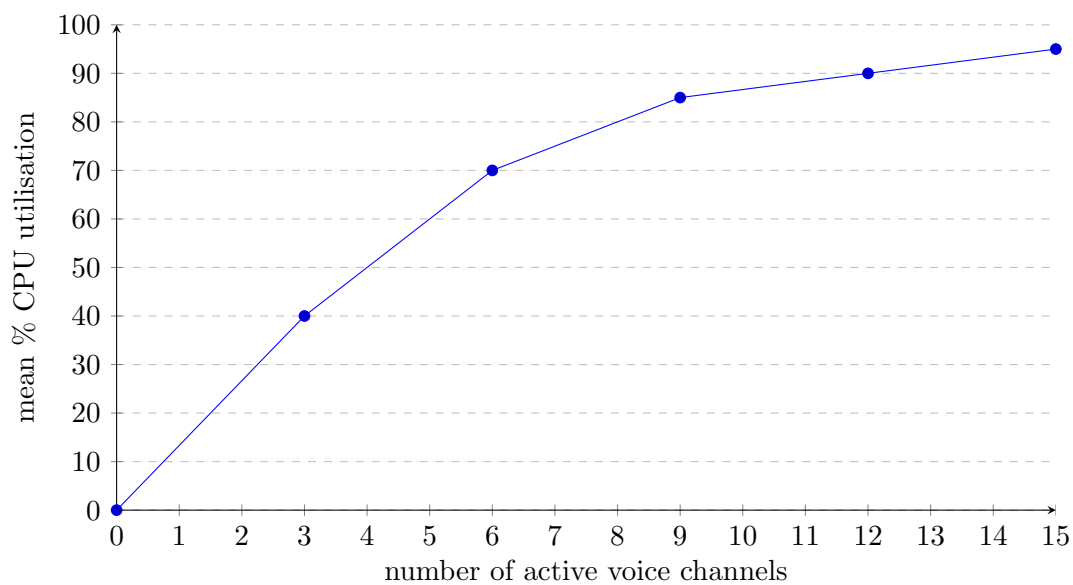


FIGURE A.1: Mean CPU utilisation given the number of active voice channels.

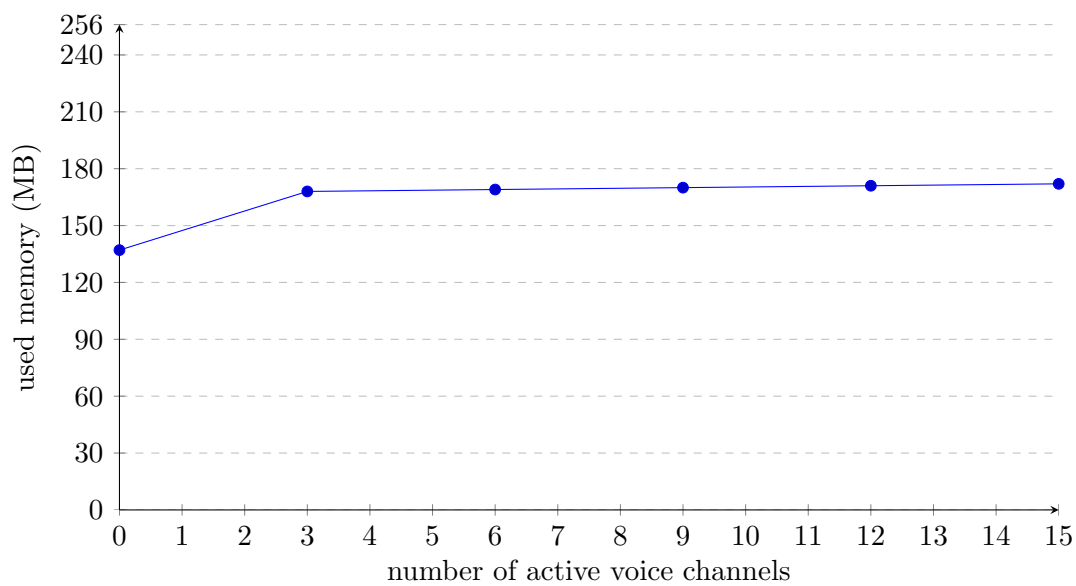


FIGURE A.2: Memory usage with active voice channels.



Sampling rate (kHz)	ogg	mp3	wav	flac
8	1.366	1.76	1.625	1.625
16	1.519	1.132	3.094	3.076
24	1.922	1.13	4.336	4.339
32	4.149	1.13	4.974	4.968
40	4.137	1.146	4.954	4.912
48	3.897	1.151	4.966	4.908

TABLE A.1: Audio streaming PEAQ scores for the common codecs.

TTS engine	Sampling rate (kHz)	Raw MOS	MOS-LQO
Festival	3	-0.230	1.114
	4	-0.257	1.096
	8	-0.314	1.657
	16	-0.400	1.034
	24	-0.364	1.027
	32	-0.421	1.023
	40	-0.480	1.018
	48	-0.466	1.017
	56	-0.523	1.017
	64	-0.574	1.017
Speect	3	-0.213	1.090
	4	-0.238	1.114
	8	-0.291	1.048
	16	-0.360	1.033
	24	-0.316	1.028
	32	-0.191	1.022
	40	-0.312	1.027
	48	-0.225	1.015
	56	-0.259	1.014
	64	-0.323	1.014

TABLE A.2: MOS-LQO scores using PESQ method to evaluate the text to audio translation gateway.

## A.2 Configuration and code snippet

This section presents the code snippet and sample configurations used in the project.

```
[demo-user] ;SIP user ID
type=friend
host=dynamic
secret=password
context=users
;deny=0.0.0.0/0
;hosts to allow registration
permit=197.239.149.0/255.255.255.224 ;match network setting

;other SIP clients included here
```

FIGURE A.3: Basic sip configuration used to setup sip clients.

Figure A.3 shows a snippet of the SIP configuration file. Each of the SIP clients have to be defined in this configuration file. Connections to the VoIP server can be made from any SIP client running on the end-user device.

```
Channel: SIP/6002@'specify_IP_address'
CallerID: test-auto
MaxRetries: 3
RetryTime: 30
WaitTime: 30
Context: call-file-test
Extension: 10
```

FIGURE A.4: Sample *call file* used to automate calls or text messages.

Figure A.4 shows a basic *call file* which is a key component of the early warning system. A SIP session is automatically established by placing the call file in `/var/spool/asterisk/outgoing/`. The call file specifies the number to send a text message to or place an automated voice call for notification or alert purposes. This action can be set to be triggered by WSN data regarding environmental variable thresholds.

```

[default] ;useful for unknown extensions

[addcaller] ;conference call setup
exten => 1,1,Originate(SIP/otherpeer,exten,conferences,100,1)

[conferences]
exten => 100,1,ConfBridge(1234)

[users]
exten=>6001,1,Dial(SIP/demo-user,20)
;other client extensions included here

exten =>6011,1,ConfBridge(1234); conference number
exten =>6598,1,Goto(demo-menu,s,1)
exten =>6000,1,Goto(demo-service-number,s,1) ;designated service number
exten =>6010, 1, VoiceMailMain(@demo-service-number)

exten => 123, 2, SendText( ${text-message-to-send} )
;exten => 123, 3, HangUp

[call-file-test]
exten => 10,1,Answer()
exten => 10,n,Wait(1)
exten => 10,n,Playback(${message-to-play-back})
exten => 10,n,Wait(1)
exten => 10,n,Hangup()

;+++++
; Background () plays in the background and is interruptible, a DTMF may be serviced at any time.
; Playback() cannot be interrupted. A user response may only be serviced once playback is complete.
;+++++
; Depending on the key pressed, pull text from storage such as the database and pass that to TTS for
; synthesis into speech. INFO1 is a function defined in the OPTION_ONE context hence the
; formation OPTION_ONE_INFO1() and so forth. The sql statements behind these functions are
; defined in /etc/asterisk/func_odbc.conf
;+++++

exten => 1,1,Playback(you-entered) ;playback key pressed
    same => n, Festival(${OPTION_ONE_INFO1()})
    same => n,Goto(s,loop)
exten => 2,1,Playback(you-entered)
    same => n, Festival(${OPTION_TWO_INFO2()})
    same => n,Goto(s,loop)
exten => 3,1,Playback(you-entered)
    same => n, Festival(${OPTION_THREE_INFO3()})
    same => n,Goto(s,loop)
exten => i,1,Playback(option-is-invalid)
    same => n,Goto(s,loop)

[demo-service-number]
exten => s,1,Answer(500)
same => n(loop), Background(welcome&thanks-for-calling-today&you-have-these-options)
same => n,WaitExten(5) ;wait 5 seconds for response
    ;menu options included here
    ;include exit option
    ;actions to take included
exten => 1,1,VoiceMail(6000@demo-service-number,u) ;leave message after pressing 1
    same => 9, HangUp() ;include key to end message posting

```

FIGURE A.5: Basic extension configuration, determines what to do with a call when received at the PBX.

```

# ***** #
# First time it's run, this script looks at text files in a given directory, uses speect TTS engine to #
# automatically map the textual representation into speech and saves the resulting audio files for #
# playback. It then monitors the text files and creates updated equivalent audio files if there are #
# any changes in the text files. #
# ***** #

import speect
import speect.audio
import speect.audio_riff
import pyinotify

voice_config_path = "/dir_path/speect_package/cmu_arctic_slt/voice.json"
audio_save_dir = "/dir_path/speect_package/audio/"
filedir = "/dir_path/speect_package/selectedText/"

voice = speect.SVoice(voice_config_path)

#At startup invoke TTS engine on the text files
#Keep watch on the files in the directory

def whenFileChanges(event):
    #get the filename of file that's been modified and synthense the text to speech
    textFile = event.pathname #full path
    name = event.name.split('.')[0] #just the file name, strip the extension

    with open(textFile, 'r') as myText:
        text = myText.read()
        utt = voice.synth(text)
        #get audio features
        audio = utt.features["audio"]
        #save to a WAV format
        audio.save_riff(audio_save_dir + name + '.wav')
        #Update log file to keep track of process activities

wm = pyinotify.WatchManager()
notifier = pyinotify.Notifier(wm)
#IN_MODIFY doesn't give filename straight, returns ~pathname/.goutputstream-XYZ instead.
#IN_CLOSE_WRITE returns the file name causing the event but may be triggered multiple times
#or unexpectedly. The behaviour seems to depend in part on the editor used to #modify the file
wm.add_watch(filedir, pyinotify.IN_CLOSE_WRITE, whenFileChanges)
notifier.loop()

```

FIGURE A.6: Automatic mapping of textual information to speech. The resulting audio files are then made accessible via a phone service.

# References

- [1] Muthoni Masinde and Antoine Bagula. A Framework for Predicting Droughts in Developing Countries Using Sensor Networks and Mobile Phones. In *Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*, number April, pages 390–393, New York, New York, USA, 2010. ACM. ISBN 9781605589503.
- [2] Hsu-Yang Kung, Jing-Shiuan Hua, and Chaur-Tzuhn Chen. Drought Forecast Model and Framework Using Wireless Sensor Networks. *Journal of information science and engineering*, 22(4):751–769, 2006.
- [3] Madsen Leif, Meggelen Jim Van, and Smith Jared. *Asterisk the Future of Telephony*. Oreilly, Sebastopol, 2 edition, 2007.
- [4] ITU-T. The E-model, a computational model for use in transmission planning, 1998. URL <http://www.itu.int/rec/T-REC-G.107-199812-S/en>.
- [5] Eric Brewer, Michael Demmer, Bowei Du, Melissa Ho, Matthew Kam, Sergiu Nedevschi, Joyojeet Pal, Rabin Patra, and Sonesh Surana. The Case for Technology in Developing Regions. *Computer*, 38(6):25–38, 2005.
- [6] Nicola J Bidwell, Mounia Lalmas, Gary Marsden, Bongiwe Dlutu, Senzo Ntlangano, William D Tucker, Matt Jones, Simon Robinson, and Elina Vartiainen. Please call ME.N.U.4EVER: designing for Callback in rural Africa. In *Proceedings of the Tenth International Workshop on Internationalisation of Products and Systems.*, pages 117–137, Kutching, 2011.
- [7] Zukile Roro, Carlos Rey-moreno, William D Tucker, and Masbulele Jay Siya. Socio-Economic Aspects of Voice-over-IP Technology in Rural SA. In *Southern African Telecommunication Networks and Applications Conference (SATNAC)*, pages 497–498, George, 2012. UWC.
- [8] Lars-Hendrik Roller and Leonard Waverman. Telecommunications Infrastructure and Economic Development: A Simultaneous Approach. *The American Economic*

- Review*, 91(4):909 – 923, February 2001. ISSN 0036-8075. doi: 10.1126/science.151.3712.867-a.
- [9] Aparajita Goyal. Information, Direct Access to Farmers, and Rural Market Performance in Central India. *American Economic Journal: Applied Economics*, 2(3): 22–45, 2010.
- [10] Emmanuel Forestier, Jeremy Grace, and Charles Kenny. Can Information and Communication Technologies be pro-poor? *Telecommunications Policy*, 26(11):623 – 646, December 2002. ISSN 03085961. doi: 10.1016/S0308-5961(02)00061-7.
- [11] Robert Jensen. The Digital Divide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *The quarterly journal of economics*, 122(3):879–924, 2007.
- [12] Dae-Heon Park, Beom-Jin Kang, Kyung-Ryong Cho, Chang-Sun Shin, Sung-Eon Cho, Jang-Woo Park, and Won-Mo Yang. A Study on Greenhouse Automatic Control System Based on Wireless Sensor Network. *Wireless Personal Communications*, 56(1):117–130, December 2011. ISSN 0929-6212. doi: 10.1007/s11277-009-9881-2.
- [13] Jeonghwan Hwang, Changsun Shin, and Hyun Yoe. Study on an Agricultural Environment Monitoring Server System using Wireless Sensor Networks. *Sensors (Basel, Switzerland)*, 10(12):11189 – 11211, January 2010. ISSN 1424-8220. doi: 10.3390/s101211189.
- [14] Tineke M Egyedi, Donna C Mehos, and Willem G Vree. New Perspectives on Inverse Infrastructures. In *Proceedings of 2009 second International Conference on Infrastructure Systems and Services: Developing 21st Century Infrastructure Networks*, pages 1 – 7, Chennai, 2009. IEEE. ISBN 9781424457915.
- [15] M. S. Islam and S. M. Z. Ahmed. The information needs and information-seeking behaviour of rural dwellers: A review of research. *IFLA Journal*, 38(2):137–147, May 2012. ISSN 0340-0352. doi: 10.1177/0340035212444513.
- [16] L Shanta Meitei and Th. Purnima Devi. Farmers information needs in rural Manipur: an assessment. *Annals of Library and information studies*, 56(1):35–40, 2009.
- [17] Sheetal K Agarwal, Arun Kumar, Amit Anil Nanavati, and Nitendra Rajput. Content Creation and Dissemination by-and-for Users in Rural Areas. In *Proceedings of 2009 International Conference on Information and Communication Technologies and Development (ICTD)*, pages 56 – 65, Doha, April 2009. IEEE. ISBN 978-1-4244-4662-9. doi: 10.1109/ICTD.2009.5426702.

- [18] Aditi Sharma Grover and Etienne Barnard. The Lwazi Community Communication Service: Design and Piloting of a Voice-based Information Service. In *Proceedings of the 20th International Conference Companion on World Wide Web*, pages 433–442, Hyderabad, India, 2011. ACM. ISBN 9781450306379.
- [19] Sambala Christian. Farmers Voice Radio, 2013. URL <http://www.farmervoice.org/sites/default/files/FVRMetaStoryMarch2013.pdf>.
- [20] Shawn Lawton Henry and McGee Liam. Accessibility, 2013. URL <http://www.w3.org/standards/webdesign/accessibility>.
- [21] A Botha, K Calteaux, M Herselman, and A S Grover. Mobile User Experience for Voice Services: A Theoretical Framework. In *Proceedings of the 3rd International Conference on Mobile Communication for Development*, pages 1–10, New Dehi, India, 2012. Karlstad University Press.
- [22] Kortum Philip. *HCI Beyond the GUI: design for Haptic, Speech, Olfactory and Other Nontraditional Interfaces*. Morgan Kaufmann, Burlington, 2008. ISBN 9780123740175.
- [23] Karen Calteaux, Aditi Sharma Grover, and Gerhard B Van Huyssteen. Business drivers and design choices for multilingual IVRs: A government service delivery case study. In *Proceedings of the 3rd International workshop on spoken languages technologies for under-resourced languages (SLTU)*, pages 29–35, Cape Town, 2012. CSIR.
- [24] Rabin Patra, Joyojeet Pal, Sergiu Nedeveschi, and Madelaine Plauche. Usage Models of Classroom Computing in Developing Regions. In *Proceedings of the International Conference on Information and Communication Technologies and Development*, number 0326582, pages 1–10, Bangalore, 2007. IEEE.
- [25] Danielsen Peter J. The Promise of a Voice-Enabled Web. *Computer*, 33(8):104–106, 2000.
- [26] R. Bruno, M. Conti, and Enrico Gregori. Mesh Networks: Commodity Multihop Ad Hoc Networks. *Communications Magazine, IEEE*, 43(3):123–131, 2005.
- [27] Mihail L Sichitiu. Wireless Mesh Networks: opportunities and challenges. In *Proceedings of World Wireless Congress*, pages 1–3, San Francisco, 2005. IEEE.
- [28] Abdulrahman Yarali, Babak Ahsant, and Saifur Rahman. Wireless Mesh Networking: A Key Solution for Emergency and Rural Applications. In *Proceedings of the second International Conference on Advances in Mesh Networks*,

- pages 143–149, Athens, Glyfada, June 2009. IEEE. ISBN 9780769536675. doi: 10.1109/MESH.2009.33.
- [29] Ian F. Akyildiz, Xudong Wang, and Weilin Wang. Wireless Mesh Networks: a survey. *Computer Networks*, 47(4):445–487, March 2005. ISSN 13891286. doi: 10.1016/j.comnet.2004.12.001.
- [30] Huang Jane-Hwa, Li-chun Wang, and Chung-Ju Chang. Coverage and Capacity of a Wireless Mesh Network. In *Proceedings of the 2005 International Conference on Wireless Networks, Communications and Mobile Computing*, pages 458–463, New York, New York, USA, 2005. IEEE. ISBN 0780393058.
- [31] J. Yuan, Z. Li, W. Yu, and B. Li. A Cross-Layer Optimization Framework for Multihop Multicast in Wireless Mesh Networks. *IEEE Journal on Selected Areas in Communications*, 24(11):2092–2103, November 2006. ISSN 0733-8716. doi: 10.1109/JSAC.2006.881617.
- [32] Francesco Licandro and Giovanni Schembra. Wireless Mesh Networks to Support Video Surveillance: Architecture, Protocol, and Implementation Issues. *EURASIP Journal on Wireless Communications and Networking*, 2007(1):1–13, 2007. ISSN 1687-1472. doi: 10.1155/2007/31976.
- [33] X. H. Wang, M. Iqbal, and X. Zhou. Design and implementation of a dual-radio wireless mesh network testbed for healthcare. In *Proceedings of the International Conference on Technology and Applications in Biomedicine*, pages 300 – 304, Shenzhen, May 2008. IEEE. ISBN 978-1-4244-2254-8. doi: 10.1109/ITAB.2008.4570647.
- [34] Wenqi (Wendy) Guo, William M. Healy, and Mengchu Zhou. ZigBee-wireless mesh networks for building automation and control. In *Proceedings of the International Conference on Networking, Sensing and Control (ICNSC)*, pages 731–736, Chicago, April 2010. IEEE. ISBN 978-1-4244-6450-0. doi: 10.1109/ICNSC.2010.5461566.
- [35] Y. Zhu, W. Zeng, H. Liu, Y. Guo, and S. Mathur. Supporting Video Streaming Services in Infrastructure Wireless Mesh Networks: Architecture and Protocols. In *Proceedings of the IEEE International Conference on Communications*, pages 1850–1855, Beijing, 2008. IEEE. ISBN 978-1-4244-2075-9. doi: 10.1109/ICC.2008.355.
- [36] Carlos Rey-Moreno, Zukile Roro, William D. Tucker, Masbulele Jay Siya, Nicola J. Bidwell, and Javier Simo-Reigadas. Experiences, challenges and lessons from rolling out a rural WiFi mesh network. In *Proceedings of the 3rd ACM Symposium on Computing for Development - ACM DEV '13*, pages 1–11, New York, New York, USA, 2013. ACM Press. ISBN 9781450318563. doi: 10.1145/2442882.2442897.



- [37] Madsen Leif, Meggelen Jim Van, and Bryant Russell. *Asterisk the Definitive Guide*. O'Reilly, Sebastopol, 3 edition, 2011.
- [38] Marco Zennaro, Bjorn Pehrson, and Antoine Bagula. Wireless Sensor Networks: a great opportunity for researchers in Developing Countries. In *Proceedings of WCITD2008*, pages 1–7, Pretoria, SA, 2008.
- [39] Ahsan Abdullah and Ahmed Barnawi. Identification of the type of agriculture suited for application of wireless sensor networks. *Russian Journal of Agricultural and Socio-Economic Sciences*, 12(12):19–33, 2012.
- [40] Robin Gebbers and Viacheslav I Adamchuk. Precision Agriculture and Food Security. *Science (New York, N.Y.)*, 327(5967):828–31, March 2010. ISSN 1095-9203. doi: 10.1126/science.1183899.
- [41] Million Mafuta, Marco Zennaro, Antoine Bagula, Graham Ault, Harry Gombachika, and Timothy Chadza. Successful Deployment of a Wireless Sensor Network for Precision Agriculture in Malawi. *International Journal of Distributed Sensor Networks*, 2013(4):1–13, 2013.
- [42] A. Matese, S.F. Di Gennaro, A. Zaldei, L. Genesio, and F.P. Vaccari. A Wireless Sensor Network for Precision Viticulture: The NAV System. *Computers and Electronics in Agriculture*, 69(1):51–58, November 2009. ISSN 01681699. doi: 10.1016/j.compag.2009.06.016.
- [43] Satish Dappin, M Vaidehi, G Nithya Nair, and TR Nair. Severity Prediction of Drought in A Large Geographical Area Using Distributed Wireless Sensor Networks. *CoRR*, 1001(3718):3042–3047, 2010.
- [44] Ivan Simeonov, Hristo Kilifarev, and Raycho Ilarionov. Embedded system for short-term weather forecasting. In *Proceedings of the International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing*, pages 24–26, New York, New York, USA, 2007. ACM.
- [45] Muthoni Masinde and Antoine Bagula. ITIKI: bridge between African indigenous knowledge and modern science of drought prediction. *Knowledge Management for Development Journal*, 7(3):274–290, 2011.
- [46] Marco Zennaro, Athanasios Floros, Gokhan Dogan, Tao Sun, Zhichao Cao, Chen Huang, Manzoor Bahader, Herve' Ntareme, and Antoine Bagula. On the Design of a Water Quality Wireless Sensor Network (WQWSN): An Application to Water Quality Monitoring in Malawi. In *Proceedings of the International Conference on Parallel Processing Workshops*, pages 330–336. IEEE, September 2009. ISBN 978-1-4244-4923-1. doi: 10.1109/ICPPW.2009.57.

- [47] Antoine Bagula, Marco Zennaro, Gordon Ingg, Simon Scott, and David Gascon. Ubiquitous Sensor Networking for Development (USN4D): an application to pollution monitoring. *Sensors (Basel, Switzerland)*, 12(1):391–414, January 2012. ISSN 1424-8220. doi: 10.3390/s120100391. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3279220&tool=pmcentrez&rendertype=abstract>.
- [48] Savio Dimatteo, Pan Hui, Bo Han, and Victor O.K. Li. Cellular Traffic Offloading through WiFi Networks. In *Proceedings of 2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems (MASS)*, pages 192–201, Valencia, October 2011. IEEE. ISBN 978-1-4577-1345-3. doi: 10.1109/MASS.2011.26.
- [49] Mathias Kretschmer, Christian Niephaus, Thorsten Horstmann, Karl Jonas, and Sankt Augustin. Providing Mobile Phone Access in Rural Areas via Heterogeneous Meshed Wireless Back-haul Networks. In *Proceedings of 2011 IEEE International Conference on Communications (ICC) Workshops*, pages 1–6, Kyoto, 2011. IEEE. ISBN 9781612849553.
- [50] Abhinav Anand, Veljko Pejovic, David L Johnson, and Elizabeth M Belding. VillageCell: Cost Effective Cellular Connectivity in Rural Areas. In *Proceedings of the 5th International Conference on Information and Communication Technologies and Development*, pages 180–189, New York, New York, USA, 2012. ACM. ISBN 9781450310451. doi: 10.1145/2160673.2160698.
- [51] Paul Gardner-Stephen and Swapna Palaniswamy. Serval Mesh Software-WiFi Multi Model Management. In *Proceedings of the 1st International Conference on Wireless Technologies for Humanitarian Relief - ACWR '11*, pages 71–77, New York, New York, USA, 2011. ACM Press. ISBN 9781450310116. doi: 10.1145/2185216.2185245.
- [52] Kun-chan Lan and Tsung-hsun Wu. Evaluating the perceived quality of infrastructure-less VoIP. In *Proceedings of the 2011 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, New York, New York, USA, July 2011. IEEE. ISBN 978-1-61284-348-3. doi: 10.1109/ICME.2011.6012150.
- [53] Neil Patel, Deepti Chittamuru, Anupam Jain, Paresh Dave, and Tapan S Parikh. Avaaj Otalo: A Field Study of an Interactive Voice Forum for Small Scale Farmers in Rural India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 733–742, New York, New York, USA, 2010. ACM. ISBN 9781605589299.
- [54] Kevin Collins, Stefan Mangold, and Gabriel-miro Muntean. Supporting Mobile Devices with Wireless LAN/MAN in Large Controlled Environments. 48(12):36–43, 2010.

- [55] Somani Patnaik, Emma Brunskill, and William Thies. Evaluating the Accuracy of Data Collection on Mobile Phones: A Study of Forms, SMS, and Voice. In *Proceedings of the International Conference on Information and Communication Technologies and Development (ICTD)*, pages 74–84, New York, New York, USA, 2009. IEEE.
- [56] Xiao-yi Lu, Zhen Fu, In-sook Lee, and Myong-soon Park. A Mechanism to Improve the Performance of Zone-Based Broadcasting Protocol in the Recovery Phase. *International Journal of Multimedia and Ubiquitous Engineering*, 3(2):63–76, 2008.
- [57] Wang Bang, Wang Wei, V Srinivasan, and Chain Chua Kee. Information Coverage for Wireless Sensor Networks. *IEEE Communications Letters*, 9(11):967–969, November 2005. ISSN 1089-7798. doi: 10.1109/LCOMM.2005.11002.
- [58] S Rappaport Theodore. *Wireless Communications Principles and Practice*. Prentice Hall, New Jersey, 2 edition, 2006. ISBN 0130422320.
- [59] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, pages 336–343, Los Alamitos, 1996. IEEE. ISBN 0-8186-7508-X. doi: 10.1109/VL.1996.545307.
- [60] Patrick M. Commarford. *Working Memory, Search and Signal Detection: implications for Interactive Voice Response system menu design*. Phd thesis, University of Central Florida, 2006.
- [61] Aditi Sharma Grover, Osamuyimen Stewart, David Lubensky, Krishna Hare, and Prabhuapada Jai. Designing Interactive Voice Response (IVR) Interfaces: localisation for low literacy users. In *Proceedings of Computers and Advanced Technology in Education*, page 8, US Virgin Islands, 2009. ACTA Press.
- [62] Johannes A Louw, Daniel R Van Niekerk, and Georg I Schliinz. Introducing the Speect speech synthesis platform. In *Proceedings of the Blizzard Challenge 2010 Workshop*, pages 1–6, Kyoto, 2010. ACTA Press.
- [63] Sheetal K Agarwal, Ketki Dhanesha, Anupam Jain, Abhishek Kumar, Nitendra Rajput, Kundan Srivastava, and Saurabh Srivastava. Organizational, Social and Operational Implications in Delivering ICT Solutions: A Telecom Web Case-study. In *Proceedings of 4th ACM/IEEE International Conference on Information and Communication Technologies and Development*, pages 2 – 9, New York, New York, USA, 2010. ACM. ISBN 9781450307871.
- [64] Pablo Montoro and Eduardo Casilari. A Comparative Study of VoIP Standards with Asterisk. In *Proceedings of 2009 Fourth International Conference on Digital*

- Telecommunications*, pages 1–6, New York, New York, USA, July 2009. IEEE. ISBN 978-0-7695-3695-8. doi: 10.1109/ICDT.2009.8.
- [65] Theo Zourzouvillys. An Introduction to Standards-Based VoIP: SIP, RTP, and Friends. *IEEE Internet Computing*, 14(2):69–73, 2010.
- [66] Raluca Musaloiu-Elefteri. *Practical wireless mesh networks and their applications*. Phd, The Johns Hopkins University, 2010.
- [67] Yair Amir, Claudiu Danilov, Michael Hilsdale, Raluca Musloiu-Elefteri, and Nilo Rivera. Fast Handoff for Seamless Wireless Mesh Networks. In *Proceedings of the 4th International conference on Mobile systems, applications and services - MobiSys 2006*, pages 83–95, New York, New York, USA, 2006. ACM Press. ISBN 1595931953. doi: 10.1145/1134680.1134690.
- [68] Hyun Jong Kim, Dong Hyeon Lee, Jong Min Lee, Kyoung Hee Lee, Won Lyu, and Seong Gon Choi. The QoE Evaluation Method through the QoS-QoE Correlation Model. In *2008 Fourth International Conference on Networked Computing and Advanced Information Management*, pages 719 –725, Gyeongju, September 2008. IEEE. ISBN 978-0-7695-3322-3. doi: 10.1109/NCM.2008.202.
- [69] Daniel Seither, A Konig, and Matthias Hollick. Routing Performance of Wireless Mesh Networks: A Practical Evaluation of BATMAN Advanced. In *2011 IEEE 36th Conference on Local Computer Networks (LCN)*, pages 897–904, Bonn, 2011. IEEE. ISBN 9781612849287.
- [70] Haytham Assem, David Malone, and Jonathan Dunne. Monitoring VoIP Call Quality Using Improved Simplified E-model. In *International Conference on Computing, Networking and Communications (ICNC)*, volume 1, pages 927 – 931, San Diego, 2013. IEEE.
- [71] ITU-R. Recommendation BS.1387-1: method for objective measurements of perceived audio quality, 2001.
- [72] ITU-T. Recommendation P.862: perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs, 2001.
- [73] Milos Cernak and Milan Rusko. An Evaluation of Synthetic Speech Using the PESQ Measure. In *Proceedings of the European Congress on Acoustics*, pages 2725–2728, 2005.
- [74] Sachin. Garg, Martin Kappes, and . Can I add a VoIP call? In *IEEE International Conference on Communications, 2003. ICC '03.*, volume 2, pages 779–783. IEEE, 2003. ISBN 0-7803-7802-4. doi: 10.1109/ICC.2003.1204432.

- [75] W. Wang, S.C. Liew, and V.O.K. Li. Solutions to Performance Problems in VoIP Over a 802.11 Wireless LAN. *IEEE Transactions on Vehicular Technology*, 54(1): 366–384, January 2005. ISSN 0018-9545. doi: 10.1109/TVT.2004.838890.
- [76] Takehiro Kawata, Sangho Shin, Andrea G Forte, and Henning Schulzrinne. Using Dynamic PCF to improve the capacity for VoIP traffic in IEEE 802.11 networks. In *Wireless Communications and Networking Conference*, pages 1589–1595, New Orleans, 2005. IEEE.
- [77] TIA. TSB-116-A: Telecommunications IP Telephony equipment Voice Quality Recommendations for IP Telephony, 2005. URL [https://eva.fing.edu.uy/pluginfile.php/67121/mod\\_folder/content/0/TR41.1-05-08-004-TSB-116A-Ballot-Version.pdf](https://eva.fing.edu.uy/pluginfile.php/67121/mod_folder/content/0/TR41.1-05-08-004-TSB-116A-Ballot-Version.pdf).
- [78] Seema Ajay Agarkar, K.D. Kulat, and R.V. Kshirsagar. WSN based Low Cost and Low Power EPM Design and Field Micro-Climate Analysis using Recent Embedded Controllers. *International Journal of Computer Applications*, 12(6):12–22, 2010.
- [79] JC Broome, JT English, JJ Marois, BA Latorre, and JC Aviles. Development of an infection model for Botrytis bunch rot of grapes based on wetness duration and temperature. *American Phytopathology Society*, 85(1):97 – 102, 1995.
- [80] Carlos Rey-Moreno, Zukile Roro, Masbulele Jay Siya, Javier Simo-Reigadas, Nicola J Bidwell, and William D Tucker. Towards a Sustainable Business Model for Rural Telephony. In *Proceedings of the III International Workshop on Research on ICT for Human Development*, pages 1–13, 2012.