

**Convivo Communicator: an Interface-Adaptive Voice over IP System  
for Poor Quality Networks**

by

Marco Antonio Escobedo Gonzalez Maiz

B.S., Electronic Systems Engineering  
Instituto Tecnológico y de Estudios Superiores de Monterrey  
Cuernavaca, Mexico (1998)

Submitted to the Program in Media Arts and Sciences, School of Architecture &  
Planning, in partial fulfillment of the requirements for the degree of

Master of Sciences in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2002


© Massachusetts Institute of Technology 2002. All rights reserved




Signature of Author .....

  
Program in Media Arts and Sciences  
May 17, 2002

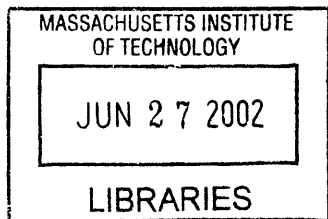
Certified by .....

  
Michael L. Best  
Research Scientist  
Thesis Supervisor

Accepted by .....

  
Andrew B. Lippman  
Chairperson

Department Committee on Graduate Students



ROTCH

# **Convivo Communicator: an Interface-Adaptive Voice over IP System for Poor Quality Networks**

by

Marco Antonio Escobedo Gonzalez Maiz

Submitted to the Program in Media Arts and Sciences, School of Architecture &  
Planning, on May 10, 2002 in Partial Fulfillment of the Requirements of the  
Degree of

Master of Science in Media Arts and Sciences

## **Abstract**

This thesis presents Convivo, a VoIP system designed to provide reliable voice communication for poor quality networks, especially those found in rural areas of the developing world. Convivo introduces an original approach to maintain voice communication interaction in presence of poor network performance: an Interface-Adaptation mechanism that suggests adjusting the application user interface to conform to one of three voice communication modalities (full duplex, half duplex, and voice messaging). The thesis proposes that changes in communication modality are an option to sustain voice communication interaction despite poor network performance. The goals of the changes in communication modality are to reduce the impact of high latency and low bandwidth on voice communication interaction, to facilitate turn taking for a high latency connection, and to sustain voice communication for extremely low bandwidth or high error links. The system was tested via a user study in Bohechio, a small village in the Dominican Republic. The study found that Interface-Adaptation helped users to maintain voice communication interaction when network performance degrades. Transitions from full duplex to voice messaging were found particularly valuable. Initial results suggest that as users get more experience with the application they would like to manually control transitions based on feedback provided by the application and their own perceived voice quality.

Thesis Supervisor: Michael L. Best

Title: Research Scientist, Program in Media Arts and Sciences

**Convivo Communicator: an Interface-Adaptive Voice over IP System  
for Poor Quality Networks**


by

Marco Antonio Escobedo Gonzalez Maiz

The following people served as readers for this thesis:

Thesis Reader .....

Christopher M. Schmandt  
Principal Research Scientist  
MIT Media Laboratory

Thesis Reader .....  


Dr. Ashok Jhunhunwala  
Professor and Head, Department of Electrical Engineering  
Indian Institute of Technology, Madras Chennai

# Acknowledgments

I am forever thankful to God for all the love He has given to me every single moment. God, thanks for this magnificent gift you gave me, thanks for all your blessings, and especially for the way you took care of me all this time. Thanks for all the strength you gave me to complete successfully this work.

I am thankful to Mike Best, my advisor, for giving me the opportunity to pursue research at the Lab. Mike, I am sincerely grateful for all your support, trust, patience, and guidance.

Thanks to Chris Schamndt and Ashok Jhunjhunwala, my readers, for their comments and interest on this project.

I would like to thank Mark Sin, Saurabh Jain, Javed Samuel, Denise Cherng, and Jae Ro who helped me develop Convivo. Thanks for all your hard work. It was a great pleasure to work with you guys.

During these two years, I was so blessed with the great friendship of Hani Shakeel, Deva Seetharam, and Leandro Burnes. You are fantastic *amigos!!* Thanks for all your support in the good and the bad times. I learned so much from all of you. Hani thanks for all the fun we shared at the lab and trips to the Dominican Republic. Deva thanks for your motivation and for being there when I most needed it. Leandro thanks for encouraging me to aim higher. The “edev” experience would not have been the same without all of you my friends. I am looking forward to see you in Mexico soon.

I would like to thank Geva Patz and Nathan Eagle, my officemates, for their good company and their support in the final countdown that insanely appeared on the screensaver of my computer. Thanks *amigos!*

I am thankful to Linda Peterson for all her help and guidance during my stay at the Media Lab.

This thesis gave me a unique opportunity to work in the Dominican Republic. Special thanks to my friends Emmanuel, Gilberto, Jhonatan, Estalin and Giovanni for the great time we had in “La Loma”, it was great hiking. Thanks to my friends Ezequiel Vicioso and Adaver Anderson for making the Convivo user study a success. Ezequiel and Adaver, remember I’ll be back to Bohechio. Thanks to all the people from Bohechio for their hospitality. Thanks also to the people from the Dominican Government: Luis Veras, Juan Garcia, and Patricia Rodriguez who helped me during my trip to the island. From other trips, I want to thank Rowland Espinoza. *Pura Vida!*

LINCOS was the platform I selected for my projects. I would like to thank all the LINCOS team, in particular to Juan Barrios and Eleonora Badilla.

Thanks to Telmex, Carlos Slim, and Javier Elguea for their initiative to invest in research and their commitment to close the “digital divide” in Mexico. Your vision opened the door for us to come to MIT and demonstrate the high quality of Mexican students.



My application to MIT would not have been a success without the help of my ITESM Professors: Dr. Ferando Ramos, Dr. Ricardo Fernandez, Dr. Hugo Zarate, and Dr. Alfredo Jacobo, and from my Softran buddies: Hebert Garcia, Patryck Valdez, and Gerardo Garcia. Thanks for your friendship and all the help and knowledge I received from you.

Thanks to all my friends and brothers from the incomparable Chapubanda. We are conquering the world as we planned many years ago!

Thanks to my friends of Pepetl, a.k.a. the “Mexican Mafia”, for all the great experiences, adventures, and good times we had together. You made this Media Lab-Boston-New England experience unique.

To Angel Alamo for his enormous friendship and prayers. Angel, thanks for always being there, in the good and bad times.

To my brothers and sisters from Congregation Leon de Juda who prayed for me all time, thank you. Special thanks to Roberto Miranda, Gregory Bishop, and Juan Carlos Huertas, you have been a great example for my spiritual life. I am deeply grateful to all my brothers and sisters from the “Brigham’s Cell Group” for your invaluable friendship and support.

I would like to thank all my brothers and sisters from Congregation Shekinah in Cuernavaca for their prayers and the spiritual support they have given me all the time. Thanks for your invaluable friendship.

I am greatly thankful to Felipe and Mari Carmen Castellanos for their guidance, love, and for sharing with me the knowledge of God. Felipe and Mari Carmen thanks for your unconditional support, and prayers. I would like to thank Marisa Castellanos for her friendship, prayers, and total support.

To my beloved Ale, my eternal gratitude. Ale, thanks for everything, for all your support, help, friendship, and love. Thanks for all the things we have shared together. Thanks because you have been a fantastic blessing to my life.

All my love to my sisters, Marisol and Ma. Alejandra, for being the best sisters anyone could ask for. *Gracias coggia y twinky!!* All my love to my parents, Marco and Ma. del Pilar, for being the best parents anyone could ask for. I love you with all my heart!! *Gracias!!*

I dedicate this thesis to my father and my mother.

*Papa y Mama:*

*Reciban todo mi amor y cariño. No puedo expresar con palabras todo lo agradecido que estoy con ustedes por la vida que han dedicado para mi, por todo su amor, paciencia y sacrificio por mi.*

*Marco*

# Contents

List of Figures .....	9
List of Tables .....	11
<b>1 Introduction</b>	<b>12</b>
1.1 Results at a Glance .....	15
1.2 Overview of Thesis .....	16
<b>2 Related Work</b>	<b>17</b>
2.1 Communication for Development .....	17
2.2 Related Systems .....	19
2.2.1 VoIP Systems .....	19
2.2.2 Robust Audio Tool .....	20
2.2.3 Community Knowledge Sharing .....	21
<b>3 Community Background</b>	<b>23</b>
3.1 Little Intelligent Communities Project .....	23
3.2 Bohechio .....	25
3.3 LINCOS in Bohechio .....	27
<b>4 System Description and Architecture</b>	<b>29</b>
4.1 Design Principles .....	29
4.2 Overall Architecture .....	31
4.3 Call Control .....	32
4.4 Message Handler .....	32
4.5 Media Streaming .....	33
4.6 QoS & Interface-Adaptation Controller .....	35
4.6.1 Internet Dynamics .....	36
4.6.2 Dynamic QoS Control General Mechanism .....	37

4.6.3	Network Status Monitor .....	38
4.6.4	QoS Controller .....	41
4.6.5	Interface-Adaptation Controller .....	43
4.6.5.1	Interface-Adaptation Principles .....	43
4.6.5.2	Communication Modalities Mapping .....	43
4.6.5.3	Network Links in Developing Countries .....	47
4.6.5.4	Interface-Adaptation Policy .....	49
4.6.5.5	Interface-Adaptation Mechanism .....	50
4.7	Graphical User Interface .....	52
4.7.1	Online Interaction .....	52
4.7.2	User Interface Areas .....	53
4.7.3	Voice Communication Modalities .....	54
4.7.4	Network Status Indicator .....	59
4.8	Directory Service .....	59
4.8.1	LDAP Services .....	60
4.8.2	Convivo Directory Service .....	61
4.8.3	Constructing Queries .....	63
<b>5</b>	<b>Evaluation</b> .....	<b>67</b>
5.1	Evaluation Objectives .....	67
5.2	Evaluation Methodology Fundamentals .....	68
5.3	Bohechio User Study .....	70
5.3.1	Overview .....	70
5.3.2	Experiment Testbed .....	71
5.3.3	Experiment Format .....	72
5.4	User Study Results .....	73
5.4.1	Sample Characteristics .....	74
5.4.2	Use of Telephone .....	74
5.4.3	Use of Computers .....	78
5.5	Statistical Report and Analysis .....	79
5.6	Experiencing Convivo .....	80

5.6.1	Network Dynamics Emulated .....	80
5.6.2	Participants Perceived Quality .....	83
5.6.3	Interface-Adaptation Mechanism .....	83
5.6.4	Communication Modalities .....	88
5.7	Graphical User Interface Perceptions .....	90
5.7.1	Conversational Graphical Representation .....	91
5.7.2	Network Status Indicator Ideas .....	91
5.7.3	Participatory Design Session .....	92
5.8	Directory Service .....	94
<b>6</b>	<b>Conclusion</b> .....	<b>96</b>
6.1	Future Work .....	97
	<b>References</b> .....	<b>100</b>

# List of Figures

Figure 1. RAT user interface .....	20
Figure 2. CKS user interface .....	22
Figure 3. LINCOS Bohechio .....	24
Figure 4. Map of the Dominican Republic .....	25
Figure 5. Convivo peer-to-peer design .....	30
Figure 6. Convivo client architecture .....	31
Figure 7. VoIP data flow .....	34
Figure 8. Dynamic QoS and Interface-Adaptation flow chart .....	37
Figure 9. Factors that contribute to the end-to-end latency .....	38
Figure 10. QoS mapping for unidirectional Internet delay and loss .....	44
Figure 11. Communication modalities mapping for unidirectional Internet delay and loss .	46
Figure 12. Communication modalities mapping for unidirectional Internet delay, loss and bandwidth	47
Figure 13. Mapping points for different types of network links .....	48
Figure 14. Sequence diagram of a request for a change in communication modality .....	51
Figure 15. Convivo user interface areas .....	53
Figure 16. Full duplex modality user interface .....	54
Figure 17. Half duplex modality user interface .....	55
Figure 18. Instant voice messaging modality user interface .....	57
Figure 19. Recording message .....	58
Figure 20. Sending message .....	58
Figure 21. Receiving message .....	58
Figure 22. Playing message .....	58
Figure 23. Network status indicator .....	59
Figure 24. Directory service access module user interface .....	62
Figure 25. Visual query “AND” operation .....	64
Figure 26. Visual query “AND” operation (simplified version) .....	64
Figure 27. Directory service user interface showing query results .....	66

Figure 28. Convivo experiment: Participant's are and control room .....	71
Figure 29. Experimental testbed .....	72
Figure 30. People with whom community members communicate by phone .....	75
Figure 31. Destination of telephone calls .....	75
Figure 32. Monthly expenses for telephone service .....	76
Figure 33. Ownership of PCs and/or access to Internet of people with whom participants ..	77
communicate by phone	
Figure 34. Reasons give for wishing to communicate using a computer .....	77
Figure 35. Participants' computer experience .....	78
Figure 36. Communication applications used in Bohechio .....	79
Figure 37. Network delay emulated during Convivo experiment .....	82
Figure 38. Packet loss emulated during Convivo experiment .....	82
Figure 39. Perceived variations in voice quality during Convivo experiment .....	83
Figure 40. Participants' level of satisfaction with transition mechanism .....	84
Figure 41. Evaluation of the degree of helpfulness of transition mechanism .....	84
Figure 42. When is a change in communication modality required? .....	85
Figure 43. When network performance degrades, what should the application needs to do?	86
Figure 44. Semi-automatic vs. automatic transitions .....	87
Figure 45. Transition mechanism preference .....	88
Figure 46. Communication modality preference .....	90

# List of Tables

Table 1. Inter-Convivo clients messages protocol .....	33
Table 2. Codec bit rates, delays and complexity .....	42
Table 3. Network status indicator: representing bandwidth available .....	59
Table 4. Icons used in the directory service user interface .....	63
Table 5. Experimental testbed computers' specifications .....	72
Table 6. Bohechio's Internet connection characteristics .....	81

# Chapter 1

## Introduction

Voice communication has been shown to be a fundamental tool for the economic and social development of rural communities in developing countries [20,41]. However, telecommunication penetration is extremely low in these regions. While there is more than one phone line for every two people in OECD countries, there is just one for every fifteen in developing countries—and one for every two hundred in the least developed countries [57].

Nowadays, officials and business owners in developing countries understand the advantages of providing access to communication and information technologies. As a result, the Internet is rapidly expanding in developing countries. The expansion in these countries includes not only urban cities but also remote and isolated rural communities. The Internet may help people to attain their development goals, but it must be used as a communication process tool [42]. For instance, rural communities with access to the Internet can benefit from new types of services such as the Voice over Internet Protocol (VoIP), which has emerged as a low-cost alternative to the Public Switch Telephone Network (PSTN) voice service. Providing VoIP services in rural communities would help to satisfy one major community need, communication.

VoIP requires a minimum quality of service (QoS) from the network to sustain a rapid delivery of voice packets and maintain voice quality. The current *best-effort* Internet model does not provide a VoIP service comparable to the traditional phone service. Over the last years,



significant advances have been made to achieve VoIP systems of good quality [3,18,21,39,56]. Thus, it is possible to achieve good QoS with the favorable network realities encountered in the developed countries.

However, network connections available in rural developing areas are not so favorable for VoIP service due to their high latency and low bandwidth. For example, in rural India 85% of copper lines cannot sustain rates higher than 8 Kbps. Wireless and satellite links are commonly used to bring Internet connectivity to such remote rural regions. These satellite connections increase the delay significantly, in the range of 250-275 ms [54]. In some countries, all national traffic goes through one satellite link (e.g. the entire country of Ghana has 12 Mbps of international bandwidth, all over satellite). As a result, the bandwidth available per Internet connection in the country is limited. Achieving good QoS under these stringent network conditions is a challenge.

Limited bandwidth and delay have a negative impact on different forms of online communication [9,18]. For example, instant text messaging can be considered, in some way, an asynchronous communication modality. Users do not expect to have a real-time interaction as in a telephone call. Text messages, which represent low amount of data, are expected to arrive with some delay. Therefore, very low bandwidth and high delay will have less impact on this form of communication. The impact would be stronger in forms of communication that require a higher level of real-time interaction, such as videoconferencing or audio conferencing. These communication modalities need high bandwidth allocations and low network delay to sustain real-time interaction. Voice communication modalities are of interest to this research, in particular three modalities are considered:

- Full duplex: two persons can speak simultaneously, real time conversation similar to a telephone call
- Half duplex: only one person can speak at a time, a button needs to be pushed and held down to talk and released to stop talking, similar to communicating using “two-way” radios
- Instant Voice Messaging: similar to a text-based instant messaging but using voice instead of text, two persons are online at the same time and they converse exchanging voice messages

This thesis presents Convivo<sup>1</sup>, a VoIP system designed to provide reliable voice communication for poor quality networks, especially those found in rural areas of the developing world. Convivo introduces an original approach to maintain voice communication interaction in presence of poor network performance: an Interface-Adaptation mechanism that suggests adjusting the application user interface to conform to one of three different voice communication modalities (full duplex, half duplex, and instant voice messaging).

In this thesis, I claim that changes in communication modality are an option to sustain voice conversation despite poor network performance. Communication modalities have different network requirements. The Interface-Adaptation principle is selecting the communication modality that would be the best form of communication to be used between Convivo clients according to the current network performance. Interface-Adaptation will reduce the impact of high latency and low bandwidth on voice communication interaction, to facilitate turn taking for a high latency connection, and to sustain voice communication for extremely low bandwidth or high error links. Additionally, Convivo dynamically adapts to the performance of network by means of an application-level QoS control that maximizes bandwidth and minimizes the end-to-end delay.

Recognizing that a large number of people in developing countries are illiterate, Convivo presents an interface with graphical representations of the voice conversation. User actions are graphically represented on the interface to maintain user interaction, presence and awareness. To facilitate user location, Convivo introduces a novel graphical interface to access a directory service. This interface gives users an easy-to-use system for browsing and searching for people who are contactable through Convivo. The directory service is based on a Lightweight Directory Access Protocol (LDAP).

Convivo is evaluated by means of a user study focused on three areas. First, the study evaluates the Interface-Adaptation mechanism. The study determines if the Interface-Adaptation mechanism helps users to communicate when network performance is poor or varying. It also illustrates what communication modality users prefer given network conditions, and defines

---

<sup>1</sup> Convivo name was inspired in the Spanish word “convivir” that means to coexist, to share.

under what circumstances the user prefers to use each of the communication modalities. Second, the study evaluates the system's usability and the user interface to determine if design principles and interface features are appropriate for the target communities. The study encourages community members to collaborate on improving Convivo's user interface. The study collects their suggestions and opinions. Third, the study tests the directory service and identifies the groups of people with whom the community members want to communicate with.

In March 2002, the system was tested via a user study in Bohechio, a rural agricultural village in the Dominican Republic, where a Little Intelligent Communities (LINCOS) community telecenter exists. The evaluation methodology and instrument were based on recommendations obtained from previous research conducted at Bohechio [51,52], and from methods of evaluating user interfaces [15]. A hands-on conversational experience with Convivo gave community members the opportunity to test the application and to express their opinions about it. A collaborative design session was conducted to obtain community's design ideas and desired features. These ideas would help developers to improve the application's functionality and user interface to meet the realities and needs of users from rural communities like Bohechio.

## **1.1 Results at a glance**

The study found that Interface-Adaptation decisions helped users to maintain voice communication interaction in presence of poor network performance. The transition from full duplex to voice messaging modality was the most helpful change when network quality degrades, and a change from full to half duplex was perceived as unhelpful. The study found that users with more computer skills would like to manually control transitions based on feedback provided by the application and their own perceived voice quality. Full duplex modality is by far the most liked communication modality because it is extremely simple to use and similar to the only form of communication modality that people from Bohechio are used to, the telephone. Nevertheless, a few of the participants mentioned voice messaging as their preferred communication modality. All participants liked and understood completely the graphical representation of conversation that appears on the interface. In particular, the portrait feature was welcomed enthusiastically.

The study found that people from Bohechio are not used to consult any kind of directory or phone book. Participants showed some difficulty in creating queries that involved more than one field. Family and friends were the most demanded categories to appear on the directory interface.

## **1.2 Overview of thesis**

This thesis is organized as follows. Chapter Two provides an overview of related work to this research. The LINCOS project and Bohechio are presented in Chapter Three. A detailed description of Convivo is presented in Chapter Four, from basic architecture to the Interface-Adaptation mechanism and user interface. The evaluation methodology and the results of the user study conducted in Bohechio are presented in detail in Chapter Five. Finally, conclusions and areas of future research are presented in Chapter Six.

# Chapter 2

## Related Work

This chapter presents related work relevant to Convivo. First is a review of communication for development. Second an overview of VoIP software applications is presented. Third is a review of the Robust Audio Tool, a VoIP application. Finally, one asynchronous messaging application designed for rural communities in developing countries is presented.

### 2.1 Communication for Development

In the last decades, much attention has been focused on the gap between developed and developing nations in terms of their access to information and communication technologies (ICTs). Initially, the efforts were conducted to increase telephone penetration in developing nations. Although the number of telephones in least developed countries has increased in the last decade, new gaps are emerging in terms of access to Internet [57]. Recent initiatives that aim to close the gap have included the G8 Digital Opportunity Task Force (DOT Force), the United Nations ICT Task Force and the UNDP/Markle Foundation Digital Opportunities Initiative.

Access to telecommunication and information systems has become vital in the economic development of every country. Communication is a basic need, especially in remote isolated communities. Poor rural areas in developing countries are beginning to have access to telecommunication systems. Rural telecommunication systems can bring new information

resources and open new communication channels. These systems can be an important piece for rural economic and social development [40].

Deploying telecommunications for rural development has become a key focus of many development agencies such as the Canadian International Development (CIDA), the World Bank, the International Telecommunications Union, and the Food and Agriculture Organization of the United Nations (FAO).

The FAO has pioneered a "communication for development" approach for catalyzing Internet services for rural stakeholders, an approach that begins with the needs of people in rural and agricultural communities [11]. The FAO promotes policy and regional coordination of Internet strategy for rural development, establishes rural Internet projects, and supports efforts to liberalize telecommunication policies in developing countries. FAO also supports rural and remote infrastructure development and promotes the development of creative Internet applications and information services for rural development.

The Grameen Village Phone project in rural Bangladesh is an example of the social and economic benefits of connectivity [41]. The project involved 950 cellular Village Phones providing telephone access to more than 65,000 people. The Village Phone program represented real savings for poor rural people for individual telephone calls. The project demonstrated that rural telephone service is very profitable. Telephones in the project brought in three times as much revenue as urban cellular phones. The Grameen experience points to a potential solution for telecom operators facing the challenge of managing the last mile of rural telecom operations: link existing and successful micro-credit organizations with telecom operators to expand public call office coverage in rural areas.

New wireless and satellites technologies have emerged to connect rural areas at lower cost than before. Rural telecommunication systems now can range from simple pay telephones on the walls of villages stores, to digital wireless cellular phones and community "telecenters" that provide Internet services. These centers contribute to increase the number of computers in rural areas in developing countries.

The Internet has opened the door to new types of applications and communication services, such as as Voice over Internet Protocol (VoIP). Also known as IP telephony, VoIP enables sending real-time voice over packet-switched networks. In the last few years, a notorious trend to integrate voice and data into one single network has produced substantial interest in VoIP. There are two main reasons for this trend: to reduce the cost of maintaining two separate networks, and to support new multimedia applications. The cost of a packet-switched network for VoIP could be as much as half of a traditional circuit-switched network for voice transmission [59]. VoIP opens the door for new applications and features, amongst others multicast conferencing, caller identification, and call forwarding. Rural communities with access to Internet connections can take advantage of low cost wireless solutions and VoIP services to satisfy the need of communication in these communities.

## **2.2 Related Systems**

A large number of commercial VoIP software applications are available. This section first provides a general overview of VoIP applications. From all VoIP applications, this section presents the Robust Audio Tool, an audio tool that provides superior voice quality when network is losing packets. A small number of projects have been focused on the context of rural communities, projects such call as technology for development. This section concludes presenting Community Knowledge Sharing (CKS), an asynchronous messaging system designed to be used in the developing world.

### **2.2.1 VoIP Systems**

A large number of commercial VoIP software applications are available. Microsoft's NetMeeting, OpenPhone, Speek Freely, Yahoo Messenger, Robust Audio Tool, Free Phone, and CUSeeME are among the most popular VoIP applications. A large list of software phones and their characteristics can be found in [17]. All audio tools provide telephone-quality speech in presence of good network performance. Most of them perform quality adaptation applying different techniques trying to achieve the best quality of service possible. Codec selection, activating silence suppression, adaptive playout adjustment, redundancy mechanism, error

concealment, dynamic jitter buffers, rate control mechanisms, and echo cancellation are some of the actions that VoIP applications enable trying to improve voice quality. Nevertheless, high QoS is not always achieved.

Microsoft's and Yahoo's messengers are communication applications that allow their users to communicate by using different communication modalities. Instant text messaging is their principal modality, but they also provide the option to establish voice or video communication. Users decide when and which communication modality to use.

### 2.2.2 Robust Audio Tool

The Robust Audio Tool (RAT) is an open source audio conferencing application that offers better sound quality relative to the network conditions than most audio tools available [45]. RAT is a VoIP application that does not perform any call services like user location.

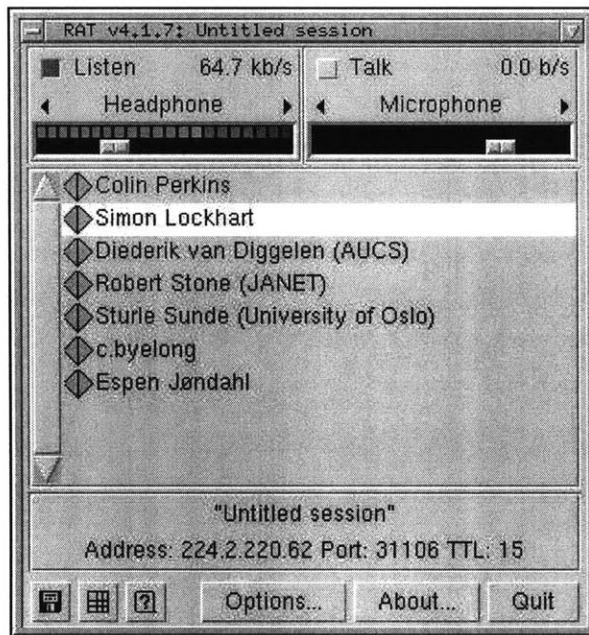


Figure 1. RAT user interface

RAT design and development attempt to address the problem of packet loss. RAT implements two sender-based repair schemes to recover from this problem: redundant transmission and



interleaving. Redundant transmission involves transmitting extra information along with the main audio packets. Redundant information is piggy-backed onto the following packet. If the original packet is lost, the redundant copy can be used in its place. The downside of this method is a reduction in sound quality but is still better than having no audio to play out in the place of the lost packet. A tradeoff exists between the amount compression used for the redundant packet and the quality of the resultant audio. Redundancy is good to provide packet loss protection for loss rates of up to 20% [45]. Interleaving involves a process where units of audio data are re-sequenced before transmission, so that originally adjacent units are separated by a guaranteed distance in the transmitted stream, and returned to their original order at the receiver. Interleaving disperses the effect of packet losses. If, for example, units are 5ms in length and packets 20ms (ie: 4 units per packet), then the first packet could contain units 1, 5, 9, 13; the second packet would contain units 2, 6, 10, 14; and so on. It can be seen that the loss of a single packet from an interleaved stream results in multiple small gaps in the reconstructed stream, as opposed to the single large gap which would occur in a non-interleaved stream.

Although interleaving does not reduce the amount of loss observed, it does significantly improve the perceived quality of an audio stream. The obvious disadvantage of interleaving is that it increases latency. This limits the use of this technique for interactive applications, although it performs well for non-interactive use. The major advantage of interleaving is that it does not increase the bandwidth requirements of a stream.

Redundant and interleaving methods, together with other features, provide a good quality of service even in presence of packet loss. However, some transient bursts are known to last up to a second, and a loss of this length cannot be repaired completely.

### **2.2.3 Community Knowledge Sharing**

Little work has been done in the areas of software development for rural communities and in software evaluation methodology in rural communities. A unique example is Community Knowledge Sharing, an Internet-enabled asynchronous messaging system designed for use in the developing world [51]. CKS implements a multi-literate design in which the system can be customized based on

the abilities and preferences of the user. CKS provides one of the first examples of an appropriate messaging system for the developing world.

CKS project study provided a detailed documentation of the methodology followed in the evaluation of CKS. Most of the recommendations and experiences gathered from that user study were taken into account for the evaluation of Convivo.

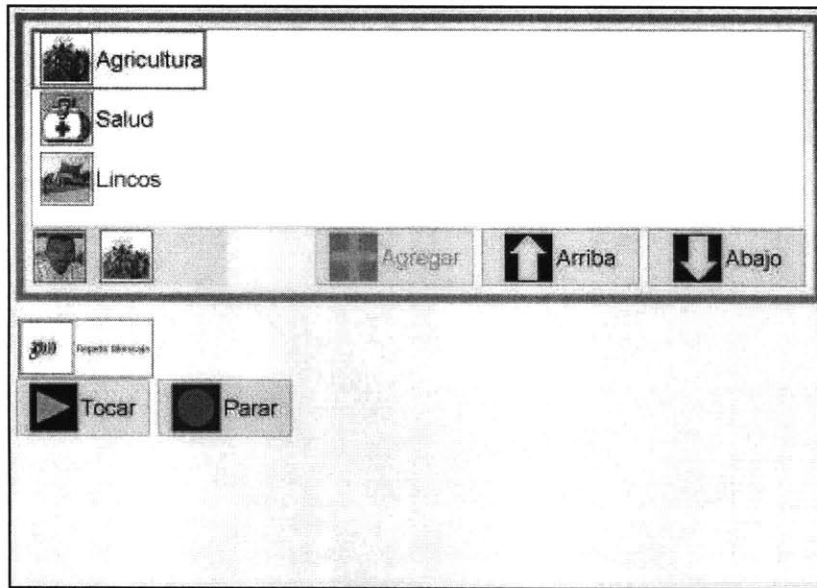


Figure 2. CKS user interface

The evaluation of CKS was conducted in Bohechio, Dominican Republic. The user study was an important source of recommendations and findings. It concluded that developers of information technologies for use in the developing work should design iconic interfaces for low literate users. Developers need to collaborate with communities and local partners to create a useful and meaningful application for the community. Moreover, the study suggests extending the multi-literate approach of CKS to a direct messaging environment to support low-cost voice and text messaging. “Doing so would go a long way towards making residents of isolated communities like Bohechio feel connected to their family in the rest of the world” states the researcher that conducted this study.

# Chapter 3

## Community Background

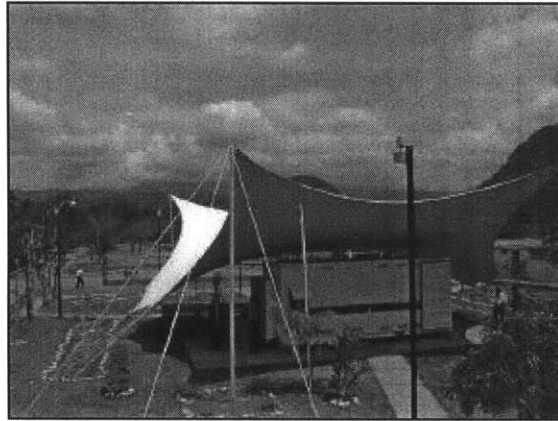
Convivo was evaluated in Bohechio, an agricultural community in the Dominican Republic. In the year 2000, the community became the first village on the island where a Little Intelligent Communities (LINCOS) telecenter was established. This section will briefly introduce the LINCOS project, its work in the Dominican Republic, and the community of Bohechio.

### 3.1 Little Intelligent Communities Project

The LINCOS project is an initiative of the Foundation for Sustainable Development of Costa Rica [31]. The Foundation, directed by former President of Costa Rica Jose Maria Figueres, promotes diverse initiatives tending to foster the development of people in countries around the world. The design and implementation of LINCOS was made possible by the collaboration of academic alliances and strategic partners: Massachusetts Institute of Technology (MIT) Media Lab, Institute of Technology of Costa Rica, University of Rochester, Microsoft, Hewlett Packard, Intel, Logo Computer System INC, Flora Foundation, Agilent Technologies, Lego, Motorola, and Avina.

LINCOS is a multi-purpose community telecenter targeted to rural isolated populations with little development opportunities. LINCOS integrates a variety of services and provides access to new communication and information technologies to empower community development. Some applications and services available at LINCOS include: telemedicine, water and soil analysis,

educative videoconference, community radio broadcasting, Internet access, fax, copy machine, scanner, digital camera, digital video camera, and several multimedia educational tools.



**Figure 3. LINCOS Bohechio**

The LINCOS structural design is based on a transportation container, selected because of its convenience, availability, and easy transportation. The container is permanently installed in a community, and an overhead canopy is added to protect it from weather conditions (Figure 3).

LINCOS communities are chosen based on an analysis of the community's needs and infrastructure. Selection process involves meeting with leaders and different community members to explain and discuss the project.

Currently, there are LINCOS sites operating in Costa Rica, the Dominican Republic, and South Africa. Other Central American governments are considering establishing LINCOS in their countries.

The LINCOS project has been a major success especially in the Dominican Republic where ten LINCOS units are in operation. The Dominican government has been very enthusiastic and committed with this project. In the year 2001, a special governmental office was created to coordinate and support all the units. Each month every LINCOS site receives RD\$25,000 pesos (approximately US\$1470 dollars) from the Dominican government. The Office also provides

basic training to the people who will manage and coordinate the telecenter in each village. Over the next two years, the government will deploy twenty more units.

### 3.2 Bohechio

Bohechio is located 220 kilometers from Santo Domingo, and 40 kilometers from the regional capital of the San Juan County in the Southern region of the Dominican Republic (Figure 4). San Juan is well known on the island because of its agricultural production of rice, coffee and beans. Bohechio is the capital of a county of the same name, which includes the communities of Arroyo Cano, El Yaque, and Montacitos. The current population of the entire Bohechio County is approximately 18,000. The town of Bohechio has approximately 572 households. The main economic activities in the village are agriculture and livestock.



Figure 4. Map of the Dominican Republic

By means of interviews with different leaders in the community, a research conducted in July 2000 revealed that Bohechio is a low developed community [52]. It is estimated that 90% of the households live below the poverty line, and 50% live in extreme poverty. The current illiteracy rate is 40%. Approximately, 60% of the population is living on one meal a day consisting of rice and eggs. That research also presents the major needs for local-economic development, which include:

- 1) Finish paving the 20-kilometer dirty road that connects Bohechio with the main east-west highways.
- 2) Improve water pump. Currently water is pumped into the town only for two hours a day.
- 3) Obtain new fields for plantation. The government expropriated productive land near Bohechio to build a hydroelectric dam.
- 4) Generate new sources of employment.
- 5) Have a stable power supply. Frequent blackouts occur in Bohechio, every evening the power level drops
- 6) Get telephone service in houses.

Nowadays, almost two years after that research field trip, the primary needs that were mentioned at that time have not been addressed by any national authority. The economic problems in the town are the same, no productive land, no sources of employment, constant migration, unstable power supply, and no paved road to Bohechio.

The road that connects Bohechio with the national east-west highway is one of the major needs. Today, only two buses provide service from Bohechio to San Juan daily. With a better road, the transportation through the town may increment. An increment in transportation would reduce the cost of transporting agricultural goods to the important national markets. Farmers would not have to wait for trucks that come to the town to buy their crops at unfair prices. Instead, they would have more opportunities to sell their crops outside Bohechio and get fairer payment for their products.

In terms of telecommunications, the community keeps requesting for private telephone lines. Prior to the installation of LINCOS, only one public phone was available. A local business was in control of that phone, charging high fees for calls. LINCOS has introduced two new telephones, representing an important competition for that local business. Thus, only these three phones are available at Bohechio. Due to the location of Bohechio on a mountainous region, cellular service does not reach the town. None of the major telephone companies that provide telecommunication services in the island has accepted the request of the people to introduce more phone lines to Bohechio.

### **3.3 LINCOS in Bohechio**

The LINCOS project has been a major milestone in the history of Bohechio. Some people consider the project to be the best thing that had happened to Bohechio at least in the past decade. Many people think that if for some reason the LINCOS closes it would be like taking one huge step back for Bohechio. Since the beginning of the project, the majority of the community members have been very enthusiastic about LINCOS.

LINCOS has had an important impact on Bohechio. Despite the economic problems that the community faces, people perceive LINCOS as an important piece of education and a key instrument for community development. Mainly, the project has contributed providing cheaper phone services, copying services, and teaching children and youth on the use of computers. Recently, adults and people from other communities close to Bohechio have started to receive classes too. This is an important milestone because this indicates that the perception of LINCOS as a project for the youth and children is changing. The radio station has been also a major success, since the majority of the community listens to the station. Educational, religious, and health programs have started to be broadcast every day.

Because Bohechio was the first community in the Dominican Republic to get a LINCOS unit, the town has received a lot of attention from international institutions, researchers, and national government. Researchers from Rochester University and MIT Media Lab have visited the community several times. In particular, MIT Media Lab's eDevelopment research group has

conducted different studies and implemented one wireless network project in the Bohechio [2,51]. In the year 2001, the Community Knowledge Sharing software application was tested in Bohechio. The wireless project, the Village Area Network, positioned Bohechio as the only LINCOS site capable to provide wireless Internet access to the entire community. Currently, three other locations in Bohechio besides LINCOS have Internet access via wireless 802.11 links: the primary school, the nun center, and the city hall. A total of twenty computers access the Village Area Network, which can provide connectivity in a radius of up to two square kilometers. At the present time, there are more computers than telephones in Bohechio. There are three phones in Bohechio, two at LINCOS and one more



# Chapter 4

## System Description and Architecture

This chapter presents a detailed description of Convivo. First, the design principles and a general overview of the system are presented. A description of the main components of the system follows, from Call Control to Interface-Adaptation and QoS controllers follows. Next, the user interface design principles and functionality are presented. This chapter concludes with the description of the directory service.

### 4.1 Design Principles

VoIP applications enable sending real-time voice over packet-switched networks. The dominant standard for transmitting multimedia in packet-switched networks is the International Telecommunication Union (ITU) Recommendation H.323 [26]. This standard uses IP/UDP/RTP encapsulation for audio. The Real-Time Protocol (RTP) is a mechanism for supporting the integration of voice, video, and data [50]. RTP headers provide sequence numbers and timestamp information needed to reassemble a real-time stream of packets. RTP makes use of the User Datagram Protocol (UDP), which is a connectionless transport layer protocol without retransmissions in case of errors. RTP has a companion protocol, the Real-Time Control Protocol (RTCP) that provides QoS feedback to the sending device, reporting on the receiver's quality of reception. The H.323 standard does not provide any QoS guarantees, but does specify that a

reliable transport protocol, such as the Transmission Control Protocol (TCP), be used for transmitting control information.

An alternative protocol for IP Telephony is the Session Initiation Protocol (SIP) [14]. Developed within the Internet Engineering Task Force (IETF), SIP enables users to initiate, negotiate and modify sessions such as an IP telephone call. Both H.323 and SIP are peer-to-peer protocols. These protocols specify how audio and video can be transported over the Internet, and what kind of signaling protocol is necessary to communicate between terminals and different entities on the network. Convivo's design emerged from the fundamentals of both protocols. From H.323, Convivo uses RTP for audio encapsulation, RTCP for QoS feedback, and the use of TCP for transmitting control. From SIP Convivo uses the call setup protocol.

The design of Convivo is based on three major principles. First, provide reliable voice communication in presence of poor network performance. Second, provide a multi-literate graphical user interface (GUI) to support users with limited literacy and computer skills, and to facilitate turn taking and maintain presence awareness given poor network performance. Third, allow users to locate each other easily.

Convivo is designed with an end-to-end (peer-to-peer) principle. Convivo clients speak to each other directly to establish a session between them. Signaling and media transmission may follow different paths. Nevertheless, Convivo clients connect to a remote server to establish presence, and locate users (Figure 5).

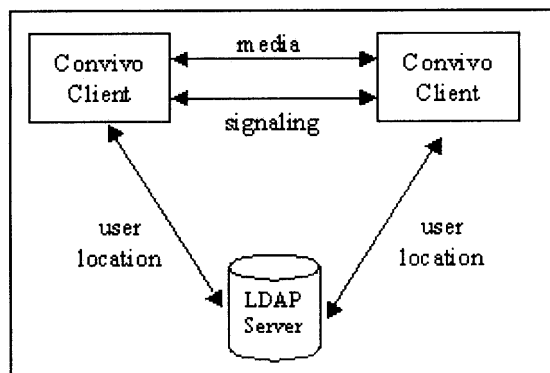


Figure 5. Convivo peer-to-peer design

## 4.2 Overall Architecture

The overall system architecture of Convivo consists of several components. The following diagram illustrates the different components and how they are interconnected to provide real-time voice communication between two parties across networks using the Internet Protocol, integrating a directory service and a multi-literate user interface.

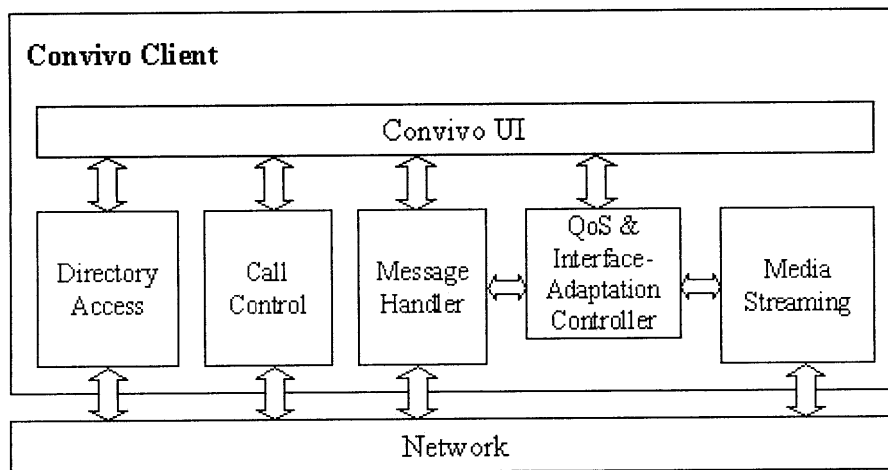


Figure 6. Convivo client architecture

Convivo client is composed of six main components (Figure 6):

1. A media streaming module which handles the audio delivery, collection and playing
2. A module which controls the Interface-Adaptation mechanism and provides dynamic QoS adaptation.
3. A message-handler module that allows Convivo clients to exchange messages
4. A call control module to allow for setup and shut down sessions
5. A directory service to allow users to find each others
6. A graphical user interface

These modules were implemented in the Java programming language. The media streaming module and the call control module are based on the Java Media Framework API package and the Java SIP implementation respectively. The other four modules are original implementations written exclusively for this thesis.

### **4.3 Call Control**

Convivo relies on the Session Initiation Protocol (SIP) to setup sessions. Basically, the call control module is based on a listener thread that runs in the background waiting for a call initiation. Upon receiving an invitation, the listener will prompt the user to accept it or reject it. Instead, if the local user starts the invitation then the listener thread will wait for the invitee's response.

Convivo clients exchange SIP messages via a TCP channel. All clients use the same port number for this purpose. Therefore, whenever the session initiation module sends an invitation it already knows the port number where the invitee is waiting for a SIP invitation.

The SIP handshake establishes the IP addresses and port at which the Convivo clients can send and receive data. Once the call invitation has been accepted, the Media and Interface-Adaptation modules are started. Then, clients initiate sending and receiving audio.

### **4.4 Message Handler**

After a Convivo session has been setup, a communication channel is established to allow the exchange of messages between clients. Convivo clients need a set of messages to control and coordinate Interface-Adaptation (communication modality transitions), turn-taking when Convivo operates in half duplex modality, and to inform the other peer about user actions. For example, a message is sent across when a user starts talking, or when a user starts or stops recording a voice message. The messages sent across are guides to update the graphical conversation section of Convivo user interface.

The Message Handler module manages this message-passing scheme, sending and receiving messages over a TCP channel. In total, five types of messages were defined to accomplish the actions mentioned in the previous paragraph (Table 1). Three messages are used to control and

coordinate Interface-Adaptation, one to control turn-taking, and one to communicate user's actions.

Command Identifier	Parameters	Description
<i>Change</i>	<i>from, to</i>	Sent when one client requests a change in communication modality. The parameters express from and to which modality change.
<i>ConfirmChange</i>	<i>from, to</i>	Sent when the client accepts a change in communication modality. The parameters express from and to which modality change.
<i>RejectChange</i>		Sent when the client rejects a change in communication modality
<i>TakeTurn</i>		Sent when the client releases the push-to-talk button in half duplex modality
<i>RemoteAction</i>	<i>action</i>	A client sends this message when one of the following actions occurs: user presses or releases the push-to-talk button, user starts or stops recording a voice message, user receives a voice message, user plays a voice message

**Table 1. Inter-Convivo clients messages protocol**

The message-handling functionality provided by this module can be easily modified and expanded to allow the exchange of extra information between Convivo clients.

## 4.5 Media Streaming

The media streaming module handles the audio delivery, collection, processing and playing. The module is based on RTP for receiving and sending voice packets over the Internet, and RTCP to provide feedback on the quality of the transmission.

To transport audio over a packet-switched network, voice is converted into a digital signal (codification process), compressed, and broken down into a series of IP packets. The packets are transported by the network, received in a playout buffer, decoded in sequential order, and played

back (Figure 7). For interactive conversation the scheme described before is used in the other direction. All real-time transport schemes use this mechanism.

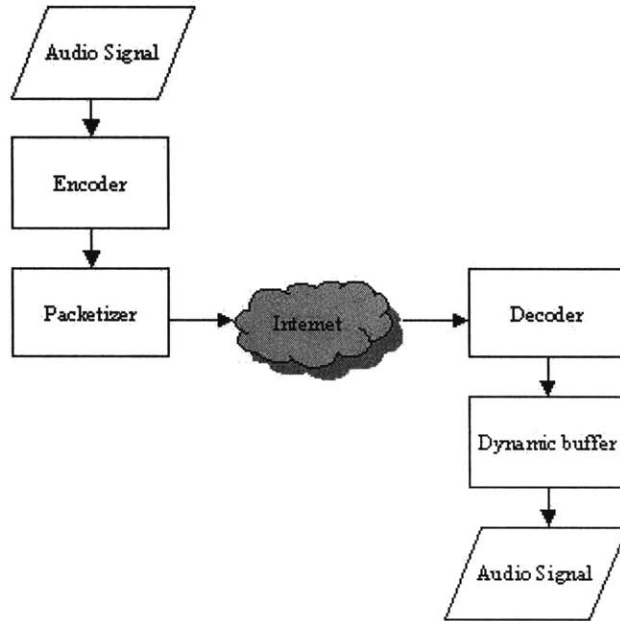


Figure 7. VoIP data flow

The media streaming module performs three main tasks:

1. Collecting audio data from the audio device, processing it, and sending RTP packets over the network; this is achieved by a Transmit thread
2. Receiving RTP packets from the network, processing them and then playing them through the audio device; this is achieved by a Receiver thread
3. Collecting RTCP reports that provide the feedback on the quality of transmission; this is also achieved by the Receiver thread

A secondary task performed by this module is processing audio to suppress periods of time when the user is not speaking. The Media Streaming module scans the audio buffer for non-silent bytes. If a silent buffer (i.e. an audio buffer is filled with silent bytes) is encountered, an internal counter is incremented. If the sequence of silent buffers is greater than a predefined limit then the module suppresses audio by discarding the output buffer. This process known as silence suppression is activated only when Convivo is operating in full duplex modality.

After the call setup process is completed, the Transmit and Receiver threads are created. Audio transmission and reception does not start until the QoS control module has performed an initial network performance analysis to select the voice transmission parameters and the voice communication modality to be used in the current session.

The process to transport audio over IP networks shown in Figure 7 has two main challenges: to maintain a short delivery time when delivering voice packets, and to maintain voice quality.

## 4.6 QoS & Interface-Adaptation Controllers

Quality of Service (QoS) is the main issue in VoIP. The current *best-effort* service available in IP networks is not sufficient to provide VoIP service comparable to the traditional Public Switch Telephony Network (PSTN) service in reliability and voice quality.

Three main influences define the QoS of VoIP: the performance of the VoIP device, the performance of the underlying network, and the end-to-end delay [39]. A particular quality level can be achieved by trading off one against the other. However, for high-quality systems the scope for trade-offs is greatly reduced. An in depth overview of VoIP QoS issues can be found in [18].

The QoS controller is responsible to achieve the main goal of Convivo: provide reliable voice communication and sustain voice communication interactions in presence of poor network performance. To achieve this goal, Convivo uses a dynamic QoS adaptation control mechanism based on two sub-controls. The first control provides reliable QoS trading off between minimizing delay and maximizing speech quality. The second control sustains voice communication interactions adjusting the user interface to conform to one of three different communication modalities: full duplex, half duplex, and instant voice messaging. The latter is referred as the Interface-Adaptation control mechanism. This section presents the details of the two sub-controls mentioned above.

### 4.6.1 Internet Dynamics

Quality of service in VoIP depends largely on the instantaneous network conditions. Understanding the end-to-end Internet dynamics is crucial for designing VoIP solutions that achieve good voice quality. This section presents a summary of Internet dynamics of importance for VoIP systems.

One of most detailed and complete analysis of end-to-end Internet dynamics characterized the prevalence of unusual network events such as out-of-order delivery and packet corruptions; found that loss events are not well-modeled independent, and that the distribution of loss events exhibits infinite variance; found that congestion periods also span a wide range of time scales [36]. Another important observation from that analysis is that asymmetric end-to-end routes are common in the Internet. Thus, the result of dividing the measured round-trip by two does not necessarily give a realistic one-way delay.

More research conducted in the area of Internet dynamics has shown that round-trip Internet delays are in the hundreds of milliseconds, and are usually correlated with packet loss [4,32,49]. In 1997, measurements made on three Internet paths in the United States during a period of six months showed that most round-trip delays are between 70 and 200 ms [30]. According to the study, observed delays are much more correlated to “hop distance” between hops than to geographical distance. The same study indicates that packet loss over these three paths is well modeled by a nonlinear Pareto distribution. This observation means that packet loss process is highly bursty, with a disproportionate amount of individual packet losses occurring in a relatively small number of bursts. Another research study made delay and loss measurements from 72 experiments over 15 different routes between hosts in Australia, Korea, USA, Sweden, France, England, and Switzerland [1]. The main findings in that study were: loss percentages were between 0.1% and 31%, long dependence in many loss and most delay series, and periodicity in both delay and loss series every 30s. Another study made measurements of delay to Stockholm from UK (33ms), US (57ms), and Argentina (273ms) [13].



## 4.6.2 Dynamic QoS Control General Mechanism

The dynamic QoS adaptation control mechanism is analogous to a feedback control system. The RTCP reports are the feedback for the QoS control module. The reports provide statistics about the current network transmission performance. The actions conducted by the module are based on the real-time statistical analysis of the following parameters: latency, bandwidth available, jitter, and packet loss rate.

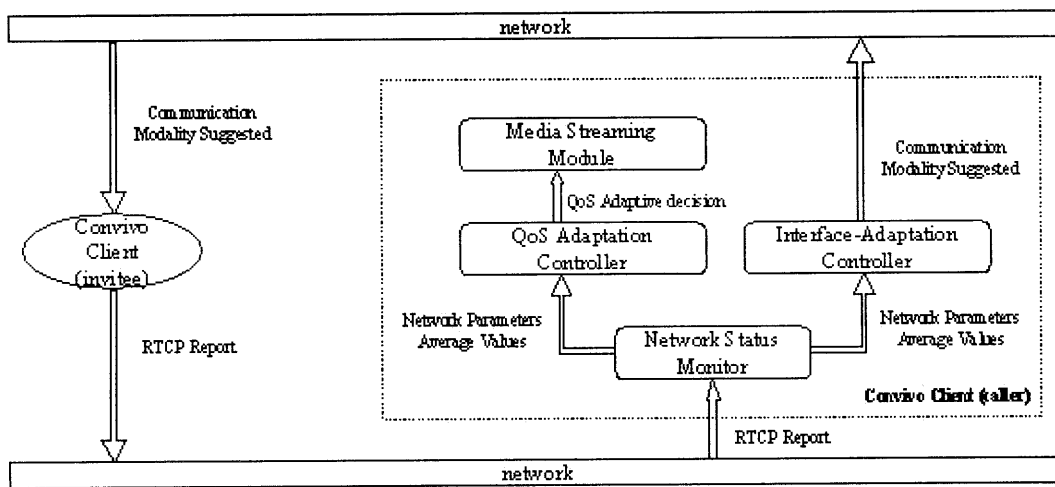


Figure 8. Dynamic QoS and Interface-Adaptation flow chart

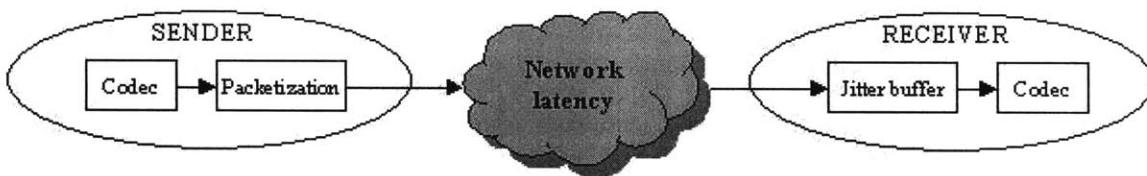
Convivo clients issue their RTCP report to the network periodically. The Network Status Monitor collects these reports and obtains the value of the network parameters (latency, bandwidth, jitter, packet loss rate). Then, the Network Status Monitor sends this data to the QoS and Interface-Adaptation controllers. These controllers make the appropriate adaptation decisions. On one hand, the QoS controller decisions have effect on the Media Streaming Module. On the other hand, the Interface-Adaptation controller decisions are sent to the other participant, if a change in communication modality is necessary both participants will experience a switch in their graphical user interface. In order to avoid two Interface-Adaptation decisions occurring simultaneously, a suggestion to switch from one communication modality to another will be made only by the Interface-Adaptation controller of the participant who initiated the session (the caller).

### 4.6.3 Network Status Monitor

VoIP systems are quite sensitive to delay and packet loss. These factors have to be measured in order to have dynamic QoS adaptation. Convivo uses the Network Status Monitor to measure the premier factors that determine voice quality: end-to-end latency, channel bandwidth available, and packet loss. Measurements are based on tests conducted by the monitor itself and RTCP reports. The values obtained are the basis for Interface-Adaptation and QoS adaptation decisions. The method used by the monitor to measure each one of the factors previously mentioned is presented below.

- **End-to-end Delay**

The end-to-end delay is the major issue in IP telephony. From coding to transport to reception to decoding, each one of these processes adds delay to the overall voice transmission end-to-end delay (Figure 9). Coding and decoding are fixed delays, while the IP network adds variable delay, and delay variance (jitter). The variable delay in the network is due to congestion and variable processing delay at routers and switches.



**Figure 9. Factors that contribute to the end-to-end latency**

The end-to-end latency is the sum of all the delays that each segment generates:

End-to-end latency = recording delay + encoding delay + compression delay (speech coder) + network delay + jitter buffer delay + decompression delay (speech coder) + decode delay (codec)

To measure network delay, the network status monitor periodically sends small UDP probe packets and measures the total time taken for a packet to go and comeback from the other peer.

The total time taken for a packet to go and comeback is the round-trip-delay. The monitor calculates the average round-trip-delay using the following equation:

$$RT = (m1 + \dots + mN) / N$$

where:

RT = average round-trip-delay

N = number of iterations

m = measured time, which is end time – start time

m1 = m for the first iteration

mN = m for Nth iteration

The average round-trip-delay value divided by two gives the average network latency.

In the Internet, packet delay is highly variable. Decompression algorithms expect incoming voice packets to arrive at a constant rate. Therefore, VoIP applications implement a jitter buffer to control the regularity in which voice packets arrive. This buffer has an impact on the delay and speech quality. The buffer holds incoming packets for a specified amount of time before forwarding them to decompression. This has the effect of smoothing packet flow, however the buffer can add significant delay. Some systems have fixed-length jitter buffers and others dynamically adjust the buffer size.

Convivo implements a jitter buffer that is adjusted dynamically upon the reception of an RTCP report. The jitter buffer depth has to be selected to balance between delay and quality. If the jitter buffer is too small, network jitter and packet loss will cause audible effects in the received voice. If the jitter buffer is too large, voice quality will be fine, but delay will be higher. The network status monitor obtains the network jitter value from the RTCP reports.

Each compression algorithm (codec) adds a fixed delay. The codec delay value is obtained from looking up the Table 2 (see Section 4.6.4). The value that corresponds to the codec currently used by the Media Stream Module is added to the total end-to-end latency.

- **Bandwidth available**

An indispensable requisite for audio transport using RTP/UDP/IP encapsulation is that enough network bandwidth be available. Therefore, based on Table 2 a bandwidth of 5.3 kbps would be enough. Actually, a total bandwidth of at least 14 Kbps is required when a speaker is talking because RTP, UDP, and IP headers add more data to voice frames. The bandwidth required when a speaker is not talking can be reduced using silence suppression, which can produce an average bandwidth saving of some 40% [39].

The network status monitor measures bandwidth by testing how fast a block of data can be sent and received. The size of the data block to send is 32 kilobytes. The monitor runs this test at the beginning of each Convivo session.

- **Packet loss rate**

IP networks do not guarantee delivery of packets. VoIP systems can support a degree of packet loss during normal operation and there can be packets dropped by the jitter buffer. Although packet loss is unavoidable, certain codecs (e.g. G.729, G.723.1, GSM-EFR) can conceal packet loss. Single packet losses have little consequences due to loss-concealment methods. In contrast, loss bursts produced by the Internet [6] can have noticeable impact on voice quality. Forward error correction (FEC) methods have been proposed to alleviate loss bursts of a small number of packets [5]. The downside of FEC methods is an increase in delay and bit rate. Because of that reason, error concealment methods are not considered in Convivo QoS actions. Instead, Convivo suggests using a different communication modality when packet loss rate is high or when constant loss bursts are present

The network status monitor calculates packet loss rate with the information provided by RTCP reports. As the RFC 1889 [50] illustrates, the packet loss rate per second can be obtained by dividing the loss fraction (observed between two RTCP reception reports) by the difference in NTP timestamps, expressed in seconds.

The RFC 1889 states: “The difference in the cumulative number of packets lost gives the number lost during that interval. The difference in the extended last sequence numbers received gives the number of packets expected during the interval. The ratio of these two is the packet loss fraction over the interval. This ratio should equal the fraction lost field if the two reports are consecutive, but otherwise not. The loss rate per second can be obtained by dividing the loss fraction by the difference in NTP timestamps, expressed in seconds. The number of packets received is the number of packets expected minus the number lost. The number of packets expected may also be used to judge the statistical validity of any loss estimates. For example, 1 out of 5 packets lost has a lower significance than 200 out of 1000.” The network status monitor follows precisely the description presented above to calculate the packet loss rate.

#### **4.6.4 QoS Controller**

The QoS controller makes proper QoS adaptive decisions that aim to maximize speech quality and minimize the end-to-end delay. Maximizing speech quality is related to speech coding, packetization efficiency, silence suppression, error concealment, jitter buffer, delay, echo, and speech level. All these factors contribute to the perceived speech quality.

- **Codec Selection**

The codec is an important component that affects voice quality. Three ITU codecs, G.723.1 [23], G.729 [24], and G.729A [25], have been designed to work well in presence of a bandwidth-, delay-, and loss-constrained environment. Table 2 lists the main characteristics of these codecs [9]. Bit rate refers to the output bit rate of the encoder. Frame size is the length of the voice signal compressed into each packet.

The lookahead delay is the amount of the next frame the coder uses to encode the current frame in order to take advantage of correlation. Frame length is the number of bytes in the encoded frames. DSP MIPS is the minimum required processor speed, in millions of instructions per second. The required RAM for each encoder is given in 16 bit-words.

Codec	6.723.1	G.729	G.729A
Bit rate (kbps)	5.3/6.4	8	8
Frame size (ms)	30	10	10
Processing delay (ms)	30	10	10
Lookahead delay (ms)	7.5	5	5
Frame length (bytes)	20/24	10	10
DSP MIPS	16	20	10.5
RAM	2200	3000	2000

**Table 2. Codec bit rates, delays, and complexity [9]**

Principally, the choice of codec in a VoIP system is dictated by the bandwidth available on the network path between call participants. The bandwidth determines the maximum bit rate for the codec, and in consequence the maximum speech quality that the system will achieve under ideal conditions. A secondary factor that influences codec selection is the latency to encode (sum of frame size, lookahead, and processing delay). From Table 2, it can be seen that while G.723.1 provides the lowest bit rate, it also adds the largest delay. G.729 encodes at higher bit rate and requires more CPU resources, but has a lower latency. Convivo selects from one of these three codec on a per-call basis to trade between bandwidth and quality. At the beginning of each call, Convivo Network Status Monitor runs a test to measure the bandwidth available and then selects the codec to be used during the call.

The QoS controller is responsible for codec selection, silence suppression activation, and jitter buffer adaptation. Codec selection is based on the bandwidth available value calculated by the network status monitor. This value determines the maximum bit rate for the codec. In general, the lower the bit rate, the lower the quality perceived by the listener. The QoS controller lookup the Table 2 and selects the codec that has a bit rate less than the actual bandwidth value.

- **Silence Suppression Activation**

QoS controller activates Silence Suppression only when Convivo clients are communicating in full duplex mode.

- **Jitter Buffer Management**

QoS Controller manages the jitter buffer storing only the required amount of buffered voice traffic. The buffer size is adapted to match the calculated network jitter.

Thus, the jitter buffer depth is set according to the RFC 1889 using the following equation [50]:

$$\text{Jitter buffer depth} = (\text{network jitter from RTCP report} * 1000) / \text{Voice Sample Rate}$$

The value obtained is the jitter buffer delay that is added to the total end-to-end latency.

#### **4.6.5 Interface-Adaptation Controller**

The Interface-Adaptation controller determines which communication modality would be the best form of communication to be used between Convivo clients according to current network performance. In order to do so, the controller analyses the data generated by the network status monitor. After the analysis is completed, the controller selects one of the three different voice communication modalities: full duplex, half duplex, and voice messaging. Each one of these forms of communication has different network requirements to provide good speech quality and conforms to different ways of dialogue. The “suggested modality” is the most appropriate form to be used to sustain voice communication interaction given the current network performance. A transition from one communication modality to another may happen only if the suggestion is different from the current communication modality, according to the “transition mechanism”.

##### **4.6.5.1 Interface-Adaptation Principles**

Various voice communication solutions can be categorized by a trade-off space. Any given solution can be characterized by its required bandwidth, end-to-end delay, and computational complexity [30]. For example, traditional telephony requires a large amount of bandwidth (64 kbps), but low computational complexity, and shows low delay. Internet audio applications, such as IP Telephony, lie in the region of low to moderate bandwidth, intermediate delay, and high computational complexity.

A nice mapping scheme from unidirectional delay and packet loss to QoS presented in [30] is shown in Figure 10. A range of QoS can be obtained based on the ITU Recommendation G.114 and codec properties. The range goes from excellent to poor quality. Excellent quality, toll quality, is for round-trip delays of less than 200 ms and low packet loss rates. The second range, good quality, is for round-trip delays from 200-300ms and higher loss rates. Limited quality is the third range, for round-trip delays from 300-800ms and high loss rates. Beyond round-trip delays of 800ms and very high loss rate, quality is extremely poor, not useful for IP Telephony. The mapping assumes a codec with good loss concealment algorithm; otherwise loss rates 5-10 percent would be enough to result in poor quality.

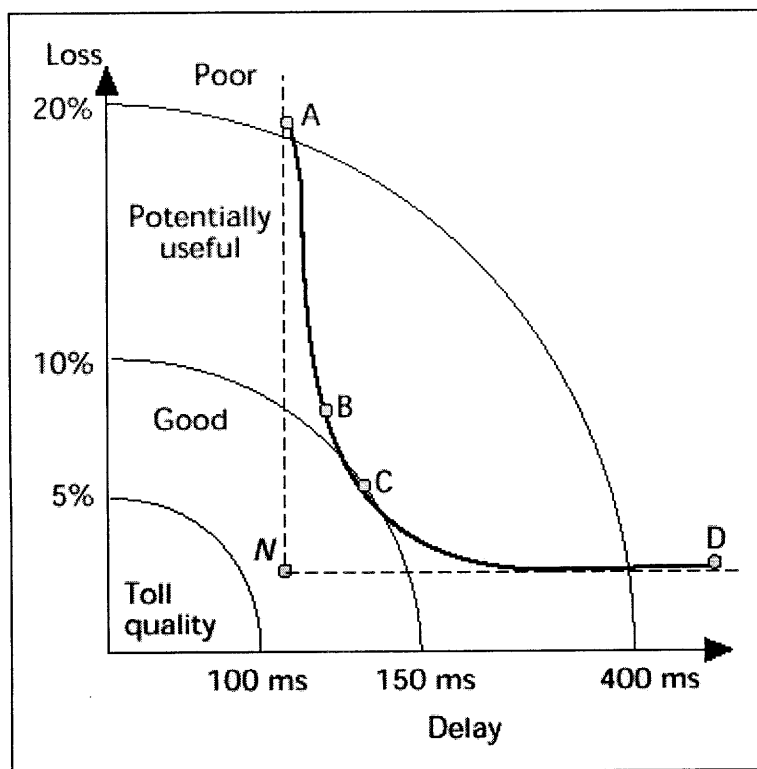


Figure 10. QoS mapping for unidirectional Internet delay and loss [30]

Point *N* in Figure 10 represents the total delay and loss that is experienced due to “network” factors, including all delays and losses up to the receiver buffer. A VoIP will never operate at this point, because codec and buffer delay also have to be considered. Point *B* represents a receiver operating under a certain amount of delay and loss. Part of this delay is due to the size of



the receiver's buffer. Increasing the buffer size will bring us to point C, which increases the delay and decreases loss rate. This decrease in loss is due to the fact that by pushing back our playout point we are not dropping as many packets. If the buffer keeps increasing toward point D and on to infinity, we reach the "network" loss rate since we no longer drop additional packets. On the other hand, if we decrease the buffer size toward point A and onward to a zero buffer length, we approach the "network" delay. However the loss rate approaches 100 percent because packets must arrive exactly at the playout point.

Available bandwidth can be added as another dimension to this mapping scheme. Bandwidth defines the threshold that determines if a synchronous real-time voice conversation is possible using VoIP. Below this threshold value (14kbps), good or poor quality of service is not the right classification, simply real-time voice interaction is not feasible. This opens the door for a new mapping scheme that relates bandwidth, delay, and packet loss to different forms of communication interaction.

#### **4.6.5.2 Communication Modalities Mapping**

Any communication modality is characterized by its required bandwidth, end-to-end delay, and packet loss. For example, videoconferencing needs high bandwidth allocation, low end-to-end delay, and moderate packet loss. In contrast, a text instant messenger demands very low bandwidth, packet loss is not an issue because TCP protocol is used for transmission, and end-to-end delay can be high because interaction is "semi-asynchronous".

A VoIP call (full duplex conversation) requires at least 14kbps in bandwidth, end-to-end delay less than 400 ms, and packet loss rate less than 5%. Most people notice delays when they exceed 250ms. The ITU-T Recommendation G.114 provides limits for one-way transmissions time, a delay of 400ms or more is unacceptable for most users [22]. To achieve the highest quality, one-way latency should be less than 150 ms. User studies indicate that telephony users find that round-trip delays greater than 300 ms feel more like a half-duplex connection than a conversation [22]. The studies showed that user tolerance of delay varies significantly from user to user. The most tolerant users were satisfied with delays of 300-800ms, while most critical

users required delays of 200 ms. Packet loss rate is a real problem when percentage of lost exceeds a certain threshold (roughly 10 %), or when loss bursts occur frequently. Even the best codecs will be unable to hide the packet loss from the user, resulting in degraded voice quality, or in worst cases making real-time voice conversation impossible.

When delay is greater than 300ms conversation conforms to a half-duplex modality. Another voice communication modality is instant voice messaging, which can be an option for voice communication when delay is extremely high (< 800 ms), when packet loss rate is notably high or when not enough bandwidth is available.

This thesis is particularly interested in three voice communication modalities: full duplex, half duplex, and instant voice messaging.

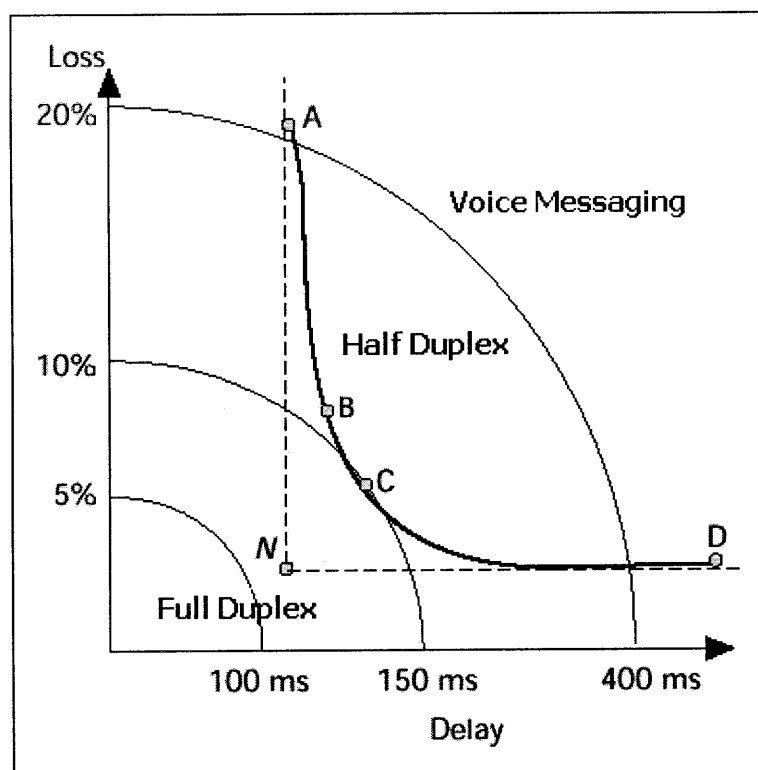


Figure 11. Communication modalities mapping for unidirectional Internet delay and loss

The analysis done for Figure 10 is valid for Figure 11, except that when we move from point to point we are also changing from communication modality. This mapping scheme is the basic idea introduced in this thesis. Convivo adapts communication modality according to the current network performance, which is named as Interface-Adaptation.

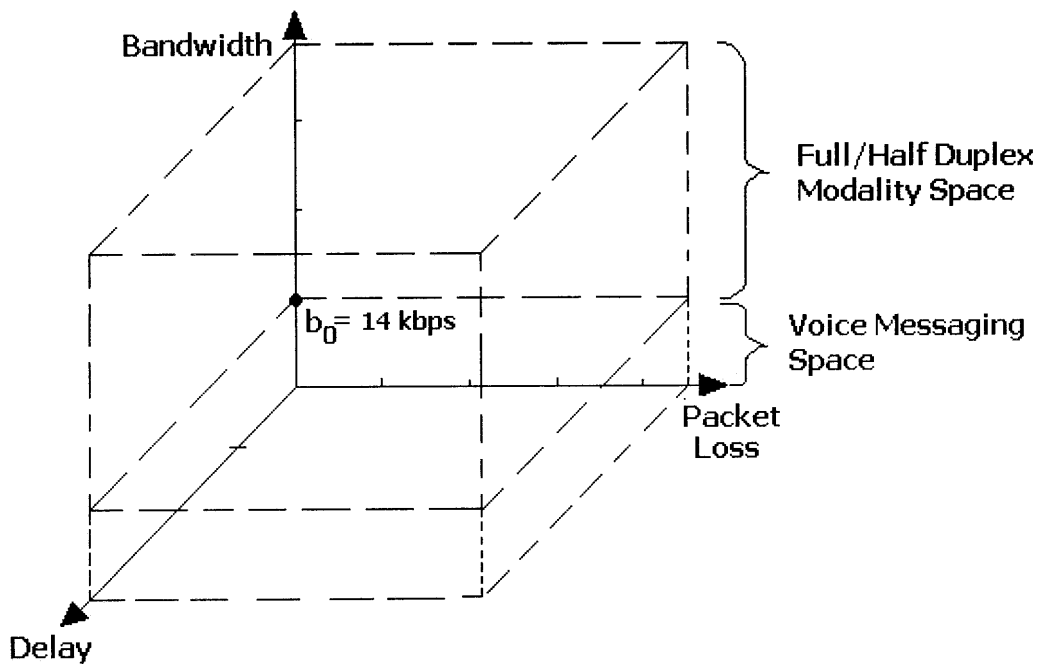


Figure 12. Communication modalities mapping for unidirectional Internet delay, loss and bandwidth

Following a similar model to the mapping presented previously, the Figure 11 shows the mapping from delay and packet loss to voice communication modalities (full duplex, half duplex, and instant voice messaging). Bandwidth available can also be added in this mapping scheme. The space is divided in two, as shown in Figure 12. The minimum bandwidth required for full or half duplex synchronous voice communication is  $b_0$ . The upper space corresponds to the full/half duplex modalities and the lower space corresponds to voice messaging modality.

#### 4.6.5.3 Network Links in Developing Countries

Given the mapping schemes presented and Internet dynamics studies we can conclude that networks, such as the ones existing in developed countries, operate in a Point  $N_0$  with

coordinates (low network delay, low network packet loss). That point  $N_0$  is either at the excellent or good quality range, with enough bandwidth to allow VoIP calls in full duplex modality.

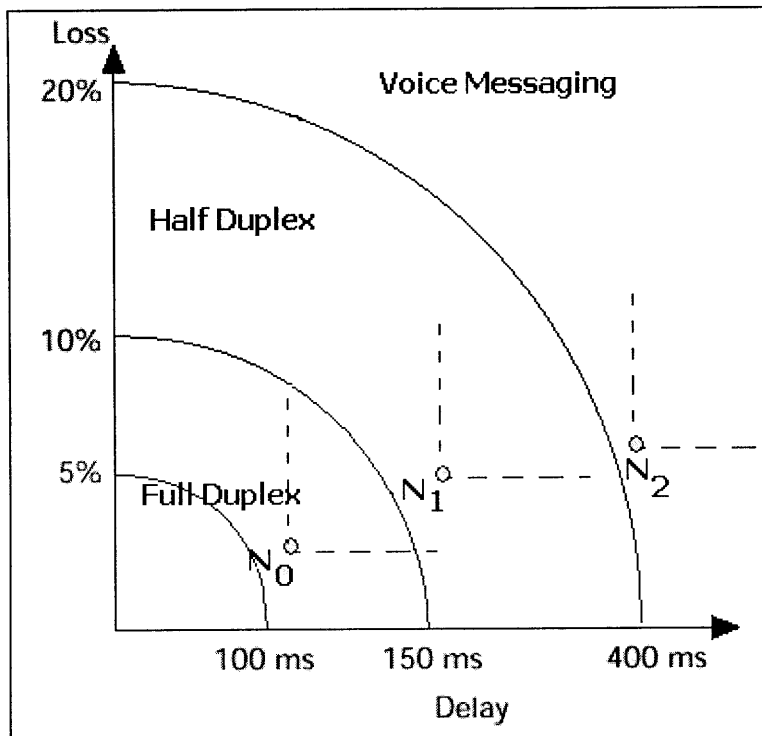


Figure 13. Mapping points for different types of network links

Point  $N_0$  might not be the same for networks in remote regions of developing countries where terrestrial communication infrastructure is scarce and based on old analog technology. In some regions, terrestrial links do not exist. Instead, wireless and satellite links are commonly used to provide Internet connectivity in developing countries. Satellite connections increase the network delay significantly, in the range of 250-275 ms, which is the one-way propagation delay on a geostationary satellite link [54]. Therefore, satellite network links in developing countries will operate at Point  $N_1$  (Figure 13) presenting higher delay than Point  $N_0$ . In some cases, more than one hop can be involved in making a connection from or to a remote. The network point at these cases would be Point  $N_2$  (Figure 13). Characteristics and issues of other types of wireless technologies that are being used to provide Internet connectivity in developing countries can be found in [48].

Satellite links are affected by sources of interference and signal power variations that manifest as errors in the transmitted information. Satellite transmission of data normally exhibits a higher error than terrestrial transmission systems such as fiber [29]. Forward error correction (FEC) can be used to minimize errors. The downside is the loss of capacity to transmit more information bits per second (i.e. FEC consumes bandwidth resources).

It presents a real challenge to provide IP telephony over satellite links with a good QoS. Little work has been done to analyze performance of VoIP over satellite systems [10,29,33,54]. The analysis, experiments, and measurements made so far indicate that it is possible to carry voice packets within the 400ms recommendation of the ITU standards. Satellite links delays can easily be higher, and experience higher packet loss rates (see Chapter 5, Bohechio satellite link characteristics).

Given the characteristics of satellite links exposed in this section and the communication modality mapping scheme, the network point of these kind of Internet links will be located very close to the half-duplex area or even at the half-duplex range if more than one satellite hop is involved.

#### **4.6.5.4 Interface-Adaptation Policy**

To determine which communication modality to be used by Convivo clients, the Interface-Adaptation controller looks at the mapping shown in figures 10 and 11. The Interface-Adaptation controller:

Suggests full duplex modality:

If (round-trip delay < 400 ms) and (bandwidth > 14 kbps) and (packet loss rate < 10%)

Suggests half duplex modality:

If (round-trip delay > 400 ms) or (bandwidth > 14kbps) and (packet loss rate < 20 %)

Suggest voice messages modality:

If (latency >> 400ms) or (14kbps < bandwidth) or (packet loss rate > 20%)

The thresholds were selected based on the ITU recommendations and research conducted in the area of Internet Dynamics.

#### **4.6.5.5 Interface-Adaptation Mechanism**

The Interface-Adaptation process follows a master-slave model. The caller is the master while the other participant is the slave. In this model, the master analyses network performance and makes suggestions, i.e. only the caller's Interface-Adaptation controller is active during the session. Caller's suggestions are sent to the slave. If a change in communication modality is necessary both participants will experience a switch in their graphical user interface. This master-slave model is applied to avoid two Interface-Adaptation suggestions made almost at the same time by two different Convivo clients, generating oscillations in the communication modality being presented to the user.

The Interface-Adaptation controller concludes what communication modality to use every thirty seconds. The value was selected based on research conducted in the field of Internet traffic dynamics [1] (see Section 4.6.1).

- **Transition Mechanism**

The transition mechanism is the scheme followed when a change in the user interface is suggested. This mechanism can operate in three different modes:

- 1) Automatic: communication modality changes in both peers automatically
- 2) Semi-automatic: Convivo prompts both users to accept or reject the transition suggested. A transition occurs if any of the users accepts the change.
- 3) Manual: users have the control to select the communication modality to be used.

When the Interface-Adaptation controller suggests a change in communication modality, the protocol shown in the next diagram is followed:

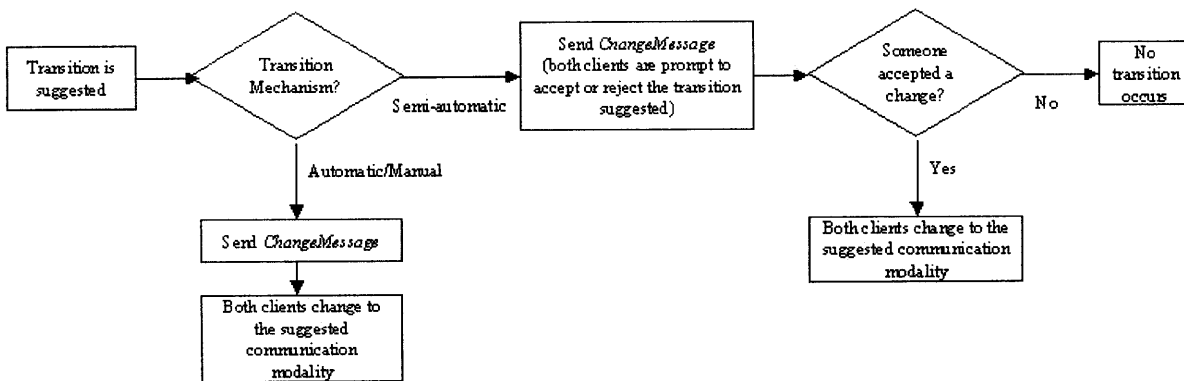


Figure 14. Sequence diagram of a request for a change in communication modality

The process starts when the master issues a *Change* message to the slave. The message contains the current modality and the new suggested modality. The slave receives the request and checks the transition mechanism mode value.

If it is set to automatic mode, the slave sends a *ConfirmChange* message to the master and switches to the new suggested modality. The master will change its user interface after receiving the confirmation message sent by the slave.

In case the mode is semi-automatic, after the master issues the *Change* message and the slave receives it, both users are prompted with a dialog box to accept or reject the transition suggested. Whoever decides first to accept or reject the transition issues a *ConfirmChange* or *RejectChange* message respectively. If change was accepted then both the master and slave switch to the new suggested modality, else they remain in the same modality.

In case of manual transition mode, the protocol followed is similar to the automatic mode. The only difference is that in automatic mode the master issues the *Change* message because the Interface-Adaptation controller suggests it, whereas in manual mode the user explicitly by means of the graphical user interface decides to change the communication modality.

## 4.7 Graphical User Interface

This section presents the user interface of Convivo. The interface was designed to handle three different voice communication modalities: full duplex, half duplex, and voice messages. The challenge was to design an interface that is simple to use to support users with a range of literacy levels and computers skills. Moreover, due to the fact that voice communication interaction is adversely affected by poor network performance, the interface has to maintain presence awareness, facilitate turn taking, and provide feedback of network status and user actions.

### 4.7.1 Online Interaction

Interaction is fundamental to human experience. Voice communication interaction will be adversely affected by poor quality of service. Limited bandwidth and high delay have a strong negative impact on online communication [28,47]. Bandwidth marks a limit of what form of communication can be used between communicants. For example, high-bandwidth connections will allow high-quality videoconferencing or high-quality voice conversations, and low-bandwidth connections will allow low-quality audio or simply text-based communication.

Transmission delays affect the way in which communicants participate in the conversation on both sides of the link. Studies conducted in 1993 found that a 570ms one-way delay in audio transmission disrupted turn-taking, with participants less likely to engage in complex interactions [55]. More studies document similar disruptions [7,35].

Karen Ruhleder and Brigitte Jordan analyzed the effect of delay in computer-mediated communication [47]. Their study concludes that delay disrupts the turn-taking system, and notes that participants cannot identify the source of trouble and repair the trouble in conversation. However, there is evidence that people pick up on the delay phenomenon and learn to moderate their manner of speaking [35]. People can become more accustomed to delay through on-going user over time. The study suggests to experiment and prototype to transform what is known about remote communication into new interfaces and applications, which can enable participants to create shared meaning and mutual understanding.



By understanding the impact of delay and bandwidth on real-time interactions, technology development can improve ways of indicating presence and attention in remote communications. The design of Convivo user interface was largely motivated by these findings and studies.

#### 4.7.2 User Interface Areas

The user interface of Convivo is divided in two main areas: the graphical conversation area, and the controls area (Figure 15).

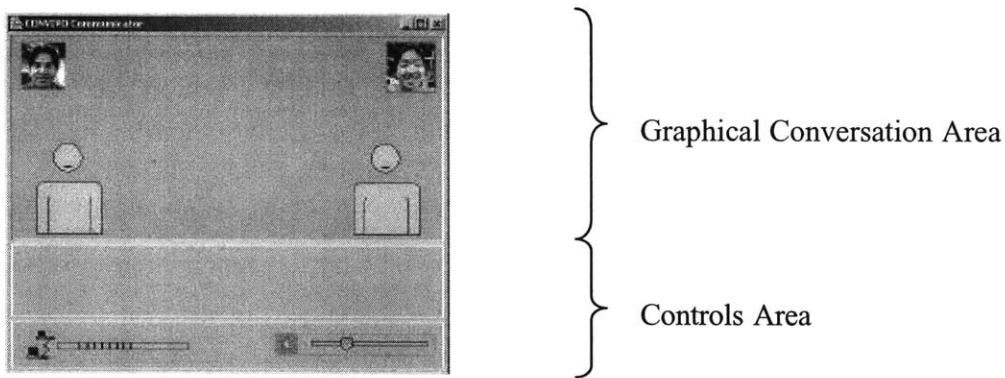


Figure 15. Convivo user interface areas

The graphical conversation area displays a visual representation of what is happening during the call. For example, the area helps users to know when the other peer is talking, listening, recording, or playing a voice message. Users are represented with an icon and a self-portrait appears on top of this icon.

From the point of view of a user, the left hand side of the graphical conversation area is related to the local user while the right area corresponds to the other peer with whom the user is communicating.

The controls area holds the buttons and controls that are required on each communication modality. Also, the volume control and the network status indicator are located in this area. The indicator is a feedback of the network status assessed by the Network Status Monitor.

### 4.7.3 Voice Communication Modalities

Convivo allow users to communicate using three different forms of voice communication: full duplex, half duplex, and instant voice messaging. The full duplex modality is similar to a telephone call where two persons can speak simultaneously. In contrast, in the half duplex modality only one person can speak at a time, a button needs to be pushed and held down to talk and released to stop talking. Half duplex modality is similar to communicating using a “walkie-talkie” radio. The instant voice messaging modality is similar to a text-based instant messaging, two persons are online at the same time and they converse exchanging voice messages. Each of these modalities follows different mechanics, and has different network requirements. The interfaces associated with the three communication modalities are shown in figures 16, 17, and 18.

- **Full Duplex**

The full duplex modality is simple, similar to a telephone call. Users can speak simultaneously and do not need to push any button, or any other control to talk. The user interface associated with this modality is shown in the following figure

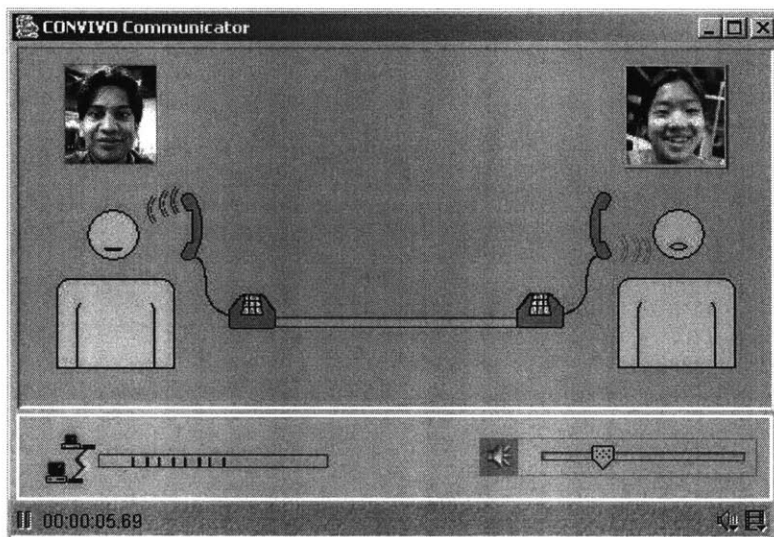


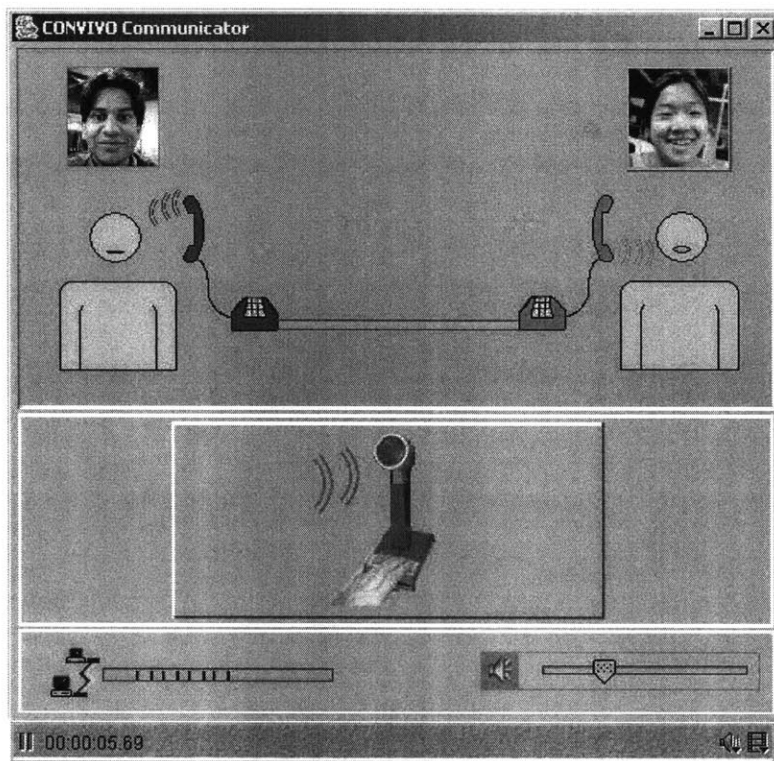
Figure 16. Full duplex modality user interface

The controls area only has the volume control and network status indicator. No other buttons are required to use the full duplex modality.

The conversation area presents both users like as if they were talking using a phone. The phones are in green. This represents that both users can speak any time. Also, when one of the participants speaks an animated icon representing voice waves is shown besides the corresponding face.

- **Half Duplex**

The half duplex modality is similar to communicating using a “walkie-talkie” (or a two-way radio). Only one person can speak at a time, a button needs to be pushed and held down to talk and released to stop talking. The user interface associated with this modality is shown in the following figure.



**Figure 17. Half duplex modality user interface**

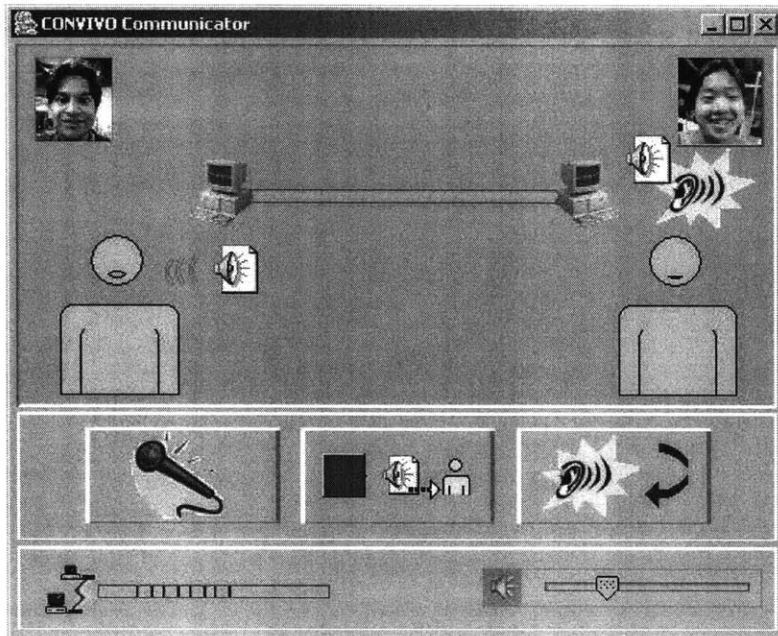
When Convivo operates in half duplex modality, a push-to-talk button appears in the controls area. When one user wants to speak this button needs to be pushed and held down to talk and released to stop talking. The push-to-talk button is enabled only on the interface of the users who holds the turn to talk.

Turn taking is controlled by a mechanism based on exchanging messages between users whenever the push-to-talk button is released. When the push-to-talk button is released, a *take turn* message is sent to the other user. Then, this user will hold the turn to talk until he has pressed and released the button again. By default, the user who started the call has the first turn to talk whenever a transition ends in the half-duplex modality.

The turn-taking scheme is reflected in the graphical conversation area. When nobody is speaking, both phone icons are grayed out. If the user who holds the turn to talk presses the push-to-talk button then the phone icon associated with him turns to green color as well as animated green voice waves appear beside his face representation. Similarly, the phone icon associated with the other user turns to red color and animated red voice waves appear besides his face representation (Figure 17).

- **Instant Voice Messaging**

The instant voice messaging modality is similar to a text-based instant messaging but using voice instead of text, two persons are online at the same time and they converse exchanging messages. The user interface associated with this modality is shown in the following figure



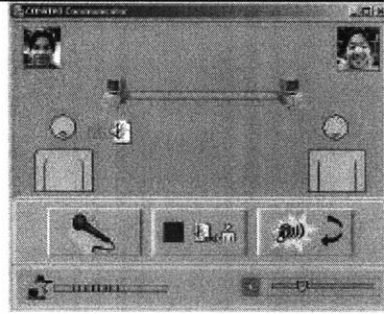
**Figure 18. Instant voice messaging modality user interface**

Three buttons are included in the controls area for this modality: StartRecording button, Stop&Send button, and PlayLastMessage button. When a user wants to compose and send a message, he has to follow two simple steps. First, he will press the StartRecording button and then start talking. Second, when he has finished talking he will push the Stop&Send button to stop recording and the message will be sent immediately to the other user. When a user receives a voice message this is played automatically, if the user wants to hear it again then he will press the PlayLastMessage button.

The sequence of recording, sending, receiving and playing a message is represented in the conversation graphical area. Animated icons will help users to know when all the mentioned actions occur. This sequence of events is shown in the following figures.

**Recording Message**

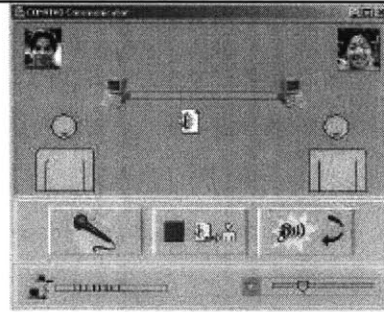
When a user starts recording, green voice waves appear in front of user's face representation and an icon of a voice message file starts to appear.



**Figure 19. Recording message**

**Sending Message**

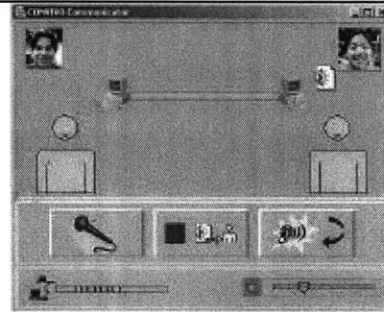
When a user finishes recording the message is sent to the other peer, the icon of a voice message moves from one peer to another indicating that the message is on its way over the network.



**Figure 20. Sending message**

**Receiving Message**

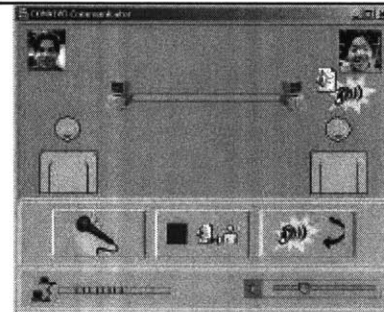
When a user receives a message, the voice message file icon that was moving from one peer to the other will be showed on top of the user who receives the message.



**Figure 21. Receiving message**

**Playing Message**

When the message received is played, the voice message file icon will also show an icon of an ear representing that the user is playing and hearing the message.



**Figure 22. Playing message**

#### 4.7.4 Network Status Indicator

Located in the left hand side of the controls area, the network status indicator gives the user a graphical feedback that reflects current network performance (Figure 23).

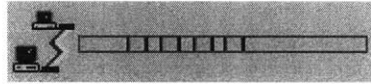


Figure 23. Network status indicator

The indicator is a horizontal bar that represents the network link between the Convivo clients participating in the call. Yellow squares, representing IP packets, traverse the bar constantly. The speed with which these squares move is relative to the current end-to-end delay. The number of squares traversing the indicator is related to the current bandwidth available (Table 3).

# of squares	Bandwidth available represented
2	bw available < 8 kbps
4	8 kpbs < bw available < 16 kbps
6	16 kbps < bw available < 32 kbps
8	bw available >> 32 kbps

Table 3. Network status indicator: representing bandwidth available

### 4.8 Directory Service

Directory services have recently emerged as an essential part of the network infrastructure. A directory is a specialized database that is read or searched more often than it is written to. A directory supports storing a wide variety of information and provides a mechanism to extend the types of information that can be stored. Directories can be centralized or distributed. They are often distributed in large scale, both in how and where information is distributed. Directories are usually replicated so that they are highly available to the clients accessing them. Replication helps increase directory performance.

## 4.8.1 LDAP Services

Directories have become standardized with the Lightweight Directory Access Protocol (LDAP) that allows applications and servers from different vendors to be developed and deployed independently [59]. Before LDAP came along, each application that needed a directory usually came built with its own directory.

New types of applications can be developed to take advantages of an online directory service. The most basic application of an online directory service is analogous to the phone book. The service provided by these kind of directories is similar to paper directories, with some differences and advantages. First, online directories can increase considerably the scale and coverage of the paper directory. Second, the directory service allows organizing the information in new ways. For example, a user can search the directory by name, phone number, or other information not contained in traditional phone books. The ability to organize information in different ways allows flexible service for searching and browsing. New kinds of searches are possible, for example, providing only part of a name the user is looking for. Also, browsing the contents of the directory can be done in different ways.

LDAP is leading the directory service space because of some important characteristics. LDAP supports a wide variety of directory-enabled applications. LDAP implementations are available for every computing platform. LDAP directories are inexpensive and easy to understand. LDAP is highly reliable and scalable.

LDAP is a “lightweight” protocol, which means that it is efficient, straightforward, and easy to implement, while still being highly functional. LDAP uses a simplified set of encoding methods and runs directly on top of TCP/IP. LDAP defines the LDAP Data Interchange Format (LDIF) a text-based format for describing directory information. LDAP version 3 is the current Proposed Internet Standard [19].

The LDAP protocol is a message-oriented protocol. The client constructs an LDAP message containing a request and sends it to the server. The server processes the request and sends the



result(s) back to the client as a series of LDAP Messages. LDAP defines nine basic protocol operations, which can be divided into three categories:

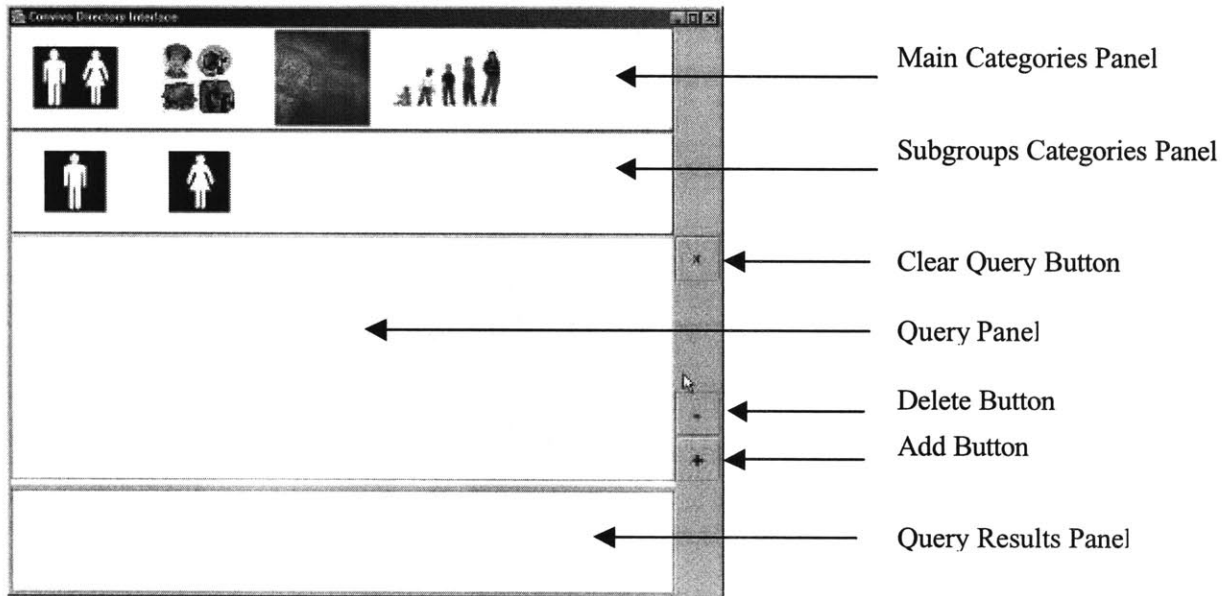
- Interrogation operations (search, compare) to ask questions of the directory
- Update operations (add, delete, modify, modify DN (rename)): to update information in the directory.
- Authentication and control operations (bind, unbind, abandon): to identify clients, to provide identification and authentication credentials, and to terminate sessions

LDAP directories can be used in VoIP applications to provide the users with the ability to locate other users who are currently on the Internet and ready to receive or make calls. Convivo uses an LDAP directory service for user location. The directory service in Convivo introduces an innovative user interface to perform interrogation operations visually and without requiring text-input.

#### **4.8.2 Convivo Directory Service**

A directory service provides Convivo users with the ability to locate other users who are currently on the Internet ready to receive or make calls. The directory stores basic information about the people registered to use Convivo. Basic information includes: name, age, community, and occupation.

Convivo users will make use of the directory service by invoking the user interface of the Directory Service Access Module. The module connects directly to the LDAP server, which is running on a remote computer, and opens up the user interface shown in Figure 24. The directory service user interface allows searching and browsing the information stored on the LDAP server.



**Figure 24. Directory service access module user interface**

The interface has four main panels and three action buttons:

- Main categories panel: shows the icons that represent the fields available for queries
- Subcategories panel: shows the icons that represent the values associated with each one of the main query fields
- Query panel: shows the visual query operation
- Search results panel: shows the results of the query
- Clear query button: clears the query and results panels to start a new query
- Add button: adds the current selected subgroup to the query panel
- Delete button: deletes the last element added to the query panel

The directory service user interface is flexible to accept a variety of main categories or fields for searching the directory service. The actual implementation shown in Figure 24 only presents four main categories: gender, age group, community, and occupation. Clicking on any icon in the main groups will change the subgroup panel to show the list of subgroups that correspond with the group. Under the main category gender, the subgroups are male, and female. The age subgroups are child, teen, adult, and elderly. Four of the LINCOS communities in the Dominican Republic are the subgroups for the community category. The options for occupations are

farming, education, health care, and religion. The icons of the four main categories and their correspondent subcategories are shown in Table 4.









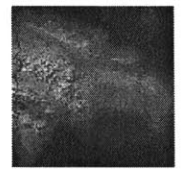



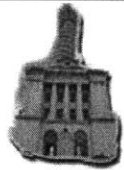





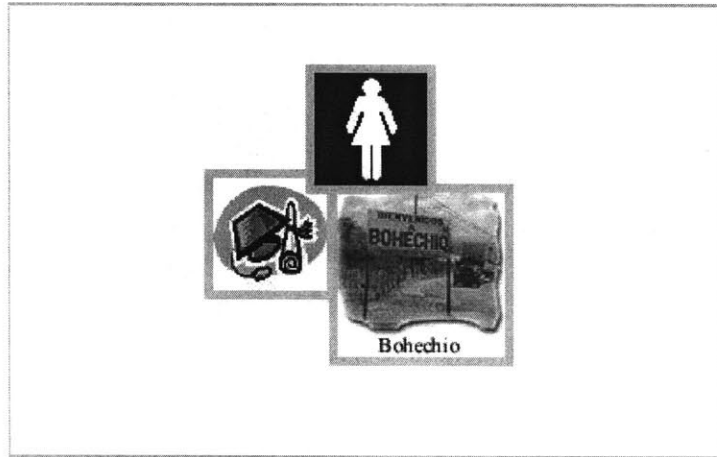
MAIN CATEGORIES		SUBCATEGORIES			
GENDER					
OCCUPATION					
COMMUNITY		 Bohechio	 El Seibo	 Oviedo	 Hato del Yaque
AGE GROUP					

Table 4. Icons used in the directory service user interface

### 4.8.3 Constructing Queries

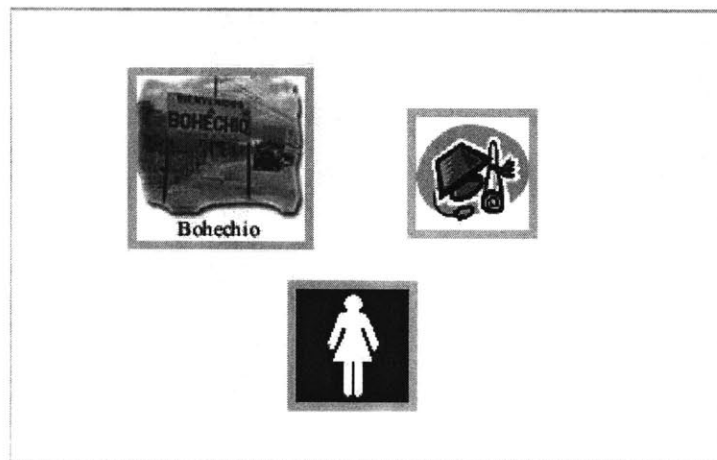
Convivo’s users will locate other users through the directory service user interface. Querying the directory server is done by means of “visual” queries. A “visual” query is constructed picking subgroups that match the desired search filter and placing them in the Query Panel. For example, to query all of the women in the town of Bohechio, the user would select the subgroup icons for females and Bohechio and add them to the Query Panel. The Query Panel is analogous to Venn diagrams. All the icons that are shown in the query panel represent a Boolean operation that is translated as the query operation sent to the LDAP server. The icons that appear in the query panel form AND and OR operations. Icons that partially overlay each other will result in the intersection of the overlapped icons, that is an AND operation. Icons or group of icons that are not intersected with any other icons will be “OR”ed into the query. It is possible to construct a

complex query involving as many icons as desired, with a mixture of “AND” and “OR” operations. For example, the query “all female students in Bohechio” is represented as shown in the Figure 25.



**Figure 25. Visual Query "AND" operation**

The query panel can be configured to allow complex Boolean operations that involve both AND and OR operations, or can be configured in a simpler way habilitating just the AND operation. In that case, the constructed query is the “AND” of all the subcategories displayed on the query panel. For example, the same query presented in Figure 25, “all female students in Bohechio” is represented as shown in Figure 26.



**Figure 26. Visual Query "AND" operation (simplified version)**

The main function performed by the directory service access module is translating the visual “query” into LDAP query operations, sending the operations to the LDAP server, retrieving and presenting results.

There are three ways to add subgroup icons to the Query Panel:

1. Double-clicking on the subgroup icon will add it to the Query Panel in a predefined position
2. Dragging the subgroup icon will add it to the Query Panel and the user can position it where he wants to drop it.
3. Clicking the “Add” button while a subgroup icon is selected will add it to the Query Panel

There are three ways to remove subgroup icons from the Query Panel:

1. Dragging the subgroup icon outside the boundaries of the Query Panel will cause it to disappear. When it disappears, it is also removed from the query.
2. Clicking the “delete” button will remove the last added subgroup icon
3. Clicking the “clear query” button will remove all subgroup icons on the Query Panel, resetting the application for a new query

To search the directory, a user would first select a group from the main searching categories panel. Once the user has selected the main group, the subgroups available for that selection will be shown. The user would browse the subgroups to construct the query. As soon as any “visual” query is constructed, the system immediately determines the appropriate search filter and connects to the LDAP server to locate matches to the query. This occurs even while the icons are dragged, making it possible to investigate different search queries quickly within the Query Panel without resetting the query. Dynamically, the directory service access module would translate the visual query shown in the Query Panel, into an LDAP query operation and send it to the LDAP server. Each time there is a change in the search filter, the results panel is updated. The results displayed are the portraits of the people who match the query (see Figure 27). The user can browse the results rolling the mouse over the portraits. To view information from any match, the user needs to click on any face in the query panel. This brings up an informational

window that covers the Query Panel. The window shows the information retrieved from the LDAP server, such as the picture, and person's attributes. If an audio message is available for the person, an icon with a speaker appears, where clicking on it will play back the audio message associated with the person, which could be an introductory description for the person. A telephone icon appears on the right of the information window. The user would decide to call the user selected, or continue browsing other results, or start a new query. To place the call, the telephone icon on the right is clicked. To start a new query the user needs to either press the clear query button or drag-and-drop all the icons off the query panel.

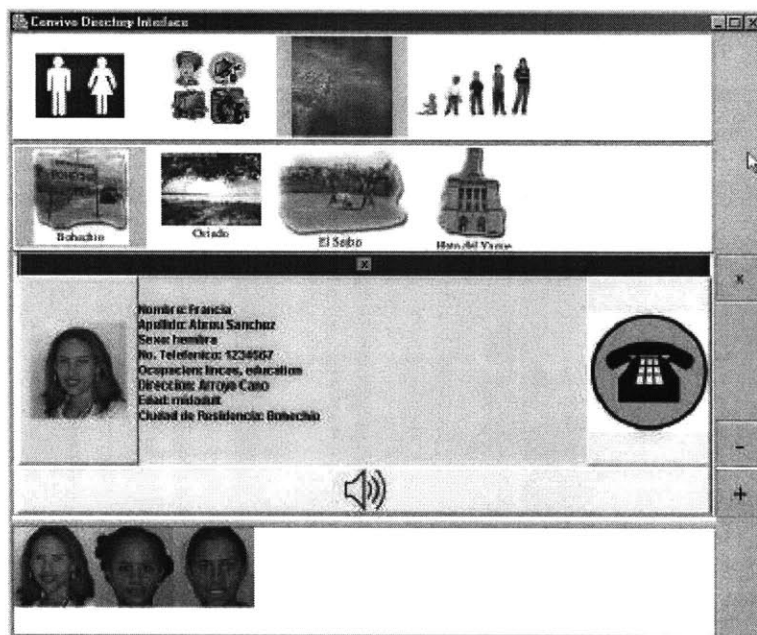


Figure 27. Directory service user interface showing query results

# Chapter 5

## Evaluation

This chapter summarizes the results of the evaluation conducted in Bohechio, Dominican Republic. It outlines the methodology used in the evaluation, describes the statistical reporting methodology, experiment format, experimental testbed, network dynamics emulated, highlights key results from participants' preferences and opinions about the Interface-Adaptation Mechanism, the communication modalities, the application graphical user interface, and the directory service.

### 5.1 Evaluation Objectives

Convivo was evaluated by means of a user study conducted in Bohechio, a rural community in the Dominican Republic. The study focused on three different areas: the Interface-Adaptation mechanism, the application graphical user interface, and the graphical directory service.

The primary research question is to determine how the Interface-Adaptation mechanism should be configured to maintain a satisfactory voice communication interaction in the presence of poor network performance. The user study will survey user's reactions, perceptions and preferences to conclude:

- 1) Does Interface-Adaptation helps users to sustain voice communication interaction when network performance varies?
- 2) How should the application react and adapt in real-time to underlying performance?
- 3) What voice communication modality should be used to improve the perceived quality for the user?
- 4) When voice communication modality should change, under what network conditions should this occur, and under what circumstances does the user prefer to use each of the communication modes?
- 5) What communication modality users prefer given network conditions?

The second area of the study focused on Convivo's graphical user interface. Convivo was designed not only to provide reliable voice communication for poor quality networks but also to make Convivo accessible to illiterate people and users with limited or no computer skills. The study in Bohechio tested the usability of Convivo, and tried to determine if the interface is understandable and appropriate for users with the previously stated characteristics. Moreover, the study closely looked at participants' communication modality preferences and interfaces usability pathologies. Furthermore, a participatory design session was conducted to gather design and feature ideas directly from locals as well as general opinions about Convivo.

The third area of this study looked at the graphical directory service. The objectives were to test the interface usability and to know what categories would make sense to include in this directory. The study investigated security and trust concerns community members have when information about personal identity is available on a public directory, which would be accessible to all Convivo's users.

## **5.2 Evaluation Methodology Fundamentals**

There is little formal literature and few documented experiences that address software evaluation in rural communities. Because of that reason, the methodology of the Convivo user study is principally based on the fields of software usability evaluation and measuring perceived quality of speech. Software usability evaluation is primarily concerned with "gathering information



about the usability or potential usability of a system in order to improve features within an interface ... or to assess a completed interface” [37]. Measuring perceived quality of speech is concerned with assessing the subjective impact of audio variables such as packet loss, delay, background noise, and echo in voice communication systems. An overview of measuring perceived quality is found in [13,39,58]. Objective [44,45] and subjective methods [27] have been designed to test end-to-end speech quality. Subjective testing is commonly used to evaluate QoS and performance of VoIP solutions.

In addition to those fields, other evaluation methodology principles and recommendations were taken from a unique example of software evaluation in rural communities, the Community Knowledge Sharing (CKS) application. CKS, a multi-literate asynchronous messaging system designed for use in the developing world, was evaluated in Bohechio. Four valuable recommendations presented were obtained from that experience and were also followed during the evaluation of Convivo. First, work with local partners to interface between the project and community. Second, use a range of data collection instruments because there is not a formal methodology for testing software in rural communities, yet is not clear what types of instruments will yield the most insight. Third, maintain flexibility to allow modifications to instruments in the field. Lastly, the community has to collaborate and participate in designing the software application.

Convivo evaluation followed methodology that reflects thinking drawn from the fields mentioned in the previous paragraphs, and is summarized below:

- Software usability evaluation was done applying the System Usability Scale (SUS) test to all participants. The SUS, which is a work of John Brooke of the digital Equipment Corporation, is a quick way to assess the overall user response to a particular interface.
- Perceived quality was measured considering standard subjective methods [27]. Participants rated the performance of the system based on a five point category scale known as the Mean Opinion Score (MOS) quality scale (in which 5=excellent, 4=good, 3=fair, 2=poor, 1=bad)

- Local partnership with LINCOS Foundation and LINCOS managers in Bohechio. They helped facilitating the research trip, and were members of the research team helping throughout the user study.
- Observation, hands-on experience, and questionnaires are useful evaluation techniques [46]. These methods were also applied during the study.
- A collaborative design session helped to get opinions of the community in order to add desired features and to improve Convivo's user interface in the future.

## 5.3 Bohechio User Study

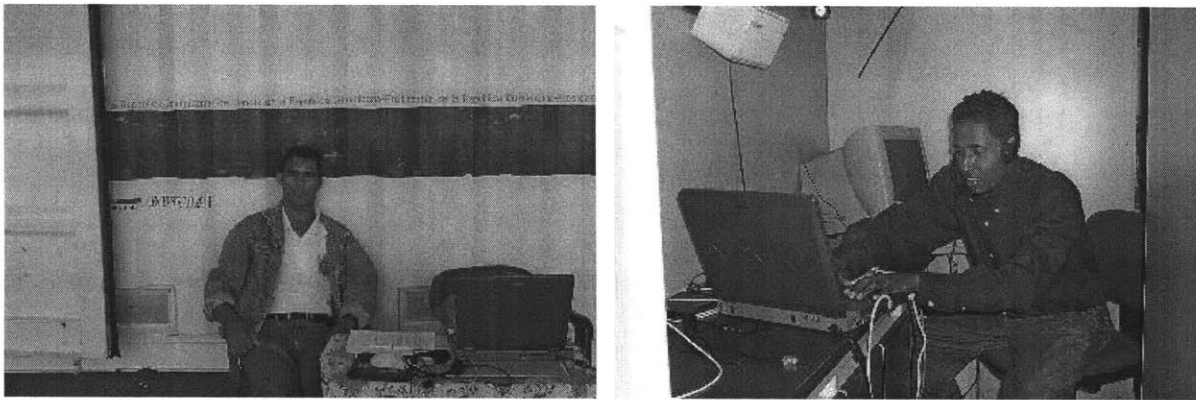
### 5.3.1 Overview

Convivo was evaluated from March 7<sup>th</sup> to 14<sup>th</sup>, 2002 in Bohechio. A research team of three people conducted this study: Marco Escobedo, Ezequiel Vicioso, and Adaver Anderson. Marco Escobedo was the leader of this evaluation. Ezequiel Vicioso and Adaver Anderson are the coordinators and managers of LINCOS in Bohechio. Their role on the team was to provide assistance to participants of the study while they were using Convivo throughout the experiment sessions. They were also assigned to observe participants' actions and comments during the experiment. The team leader gathered their observations and comments at the end of the study on separate interviews.

In total, twenty-eight people participated in this study, three helped to refine the experiment methodology and twenty-five to gather data. Only the quantitative responses gathered from those twenty-five sessions were included in the results presented in the following sections. Participants were randomly selected to get a range of age, education, and computer skills level. Additional data was collected from a participatory design session and from interviews held with the coordinators of LINCOS, Ezequiel and Adaver.

### 5.3.2 Experiment Testbed

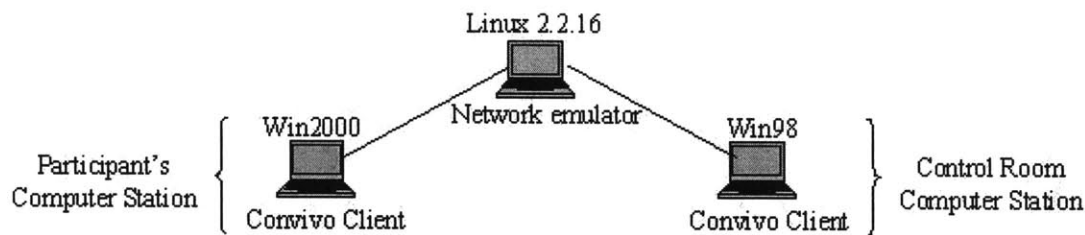
In order to evaluate Convivo, real voice communication has to be established between two different computers. The research team set two different working areas: the participant's desk station and the control room. Neither visual contact nor audio contact was possible between these areas. The participant's station was located outside the LINCOS container, while the control room was inside the telemedicine room of the LINCOS container (Figure 28).



**Figure 28. Convivo experiment: Participant's area and control room**

The experimental testbed topology used during the evaluation is shown in Figure 29. Three computers were used (see Table 5, for specific details): one computer per working area linked through a computer running Linux operating system which was set up as a router. In addition, the Linux machine was running the NIST Net network emulator tool [34].

NIST Net allows a single Linux PC set up as a router to emulate a wide variety of network conditions. The tool is designed to allow controlled, reproducible experiments with network performance sensitive/adaptive applications and control protocols in a simple laboratory setting. NIST Net can emulate the critical end-to-end performance characteristics imposed by various wide area network situations (e.g., congestion loss). NIST Net allows the emulation of complex performance scenarios, including: packet delay distributions, congestion and background loss, bandwidth limitation, and packet reordering.



**Figure 29. Experimental testbed**

The NIST Net network emulator applies selected performance “effects” to the IP packets of the stream passing through the router where it runs.

By this means, different network dynamics between the participant’s and control room computers were emulated during the experiment sessions.

	PC model and processor	Operating System	RAM
Participant’s Computer Station	Sony Vaio Laptop Pentium III – 700 Mhz	Windows 2000	196 Mbytes
Control Room’s Computer Station	HP desktop Celeron - 300 Mhz	Windows 98	64 Mbytes
Network Emulator Computer	Toshiba Tecra780 Laptop Pentium MMX – 166 Mhz	Red Hat Linux 7.0 Linux Kernel 2.2.16	64 Mbytes

**Table 5. Experimental testbed: computers’ specifications**

### 5.3.3 Experiment Format

Each experiment session was divided in five sections: 1) introduction of the experiment, 2) participant information, 3) Convivo experiment, 4) directory service experiment, and 5) interview.

The experiment session start providing the participant a description of the experiment, followed by a tutorial on how to use Convivo. The interviewer explains the difference between the voice communication modalities showing screenshots of each one of them. Then, if the participant has

no questions, the interviewer takes a picture of the participant, and uploads it to the computer. When this section of the experiment finishes, the next immediately follows.

Originally, the team leader would conduct all sections. Either Ezequiel Vicioso or Adaver Anderson would help in the third section using Convivo to communicate with the experiment participant at another computer station. After the first two trial sessions, the team leader realized that it would be better to exchange places with them during the third section. In that way, the team leader would be at the other computer station using Convivo and they would be providing assistance to participants in the other station. On one hand, this approach would give the team leader the opportunity to keep talking to participants all the time by means of Convivo. Also, he would be able to take additional notes, to be more in touch with participants, and to experiment the same communication channel and network performance as the participants. On the other hand, having either Ezequiel or Adaver working closely to the participant at the evaluation station would give the team other advantages. First, they could establish a higher level of trust with the participants. As a result, participants would feel more comfortable and more confident to use the application, to ask questions, and to ask for assistance. Furthermore, whoever was in turn for assistance would be able to observe participants' impressions and reactions directly, as well as to note the problems that participants had while using Convivo. Finally, they could potentially get comments or suggestions from participants that maybe they would not tell the team leader because he is not local. At the end of the study, this proved to be true, Ezequiel and Adaver received feedback and comments that participants did not tell the team leader during the experiment session.

## **5.4 User Study Results**

The following sections present the qualitative results of the twenty-five experiment sessions conducted in Bohechio user study. Where valuable qualitative results were gathered, they are included in the analysis.

### **5.4.1 Sample Characteristics**

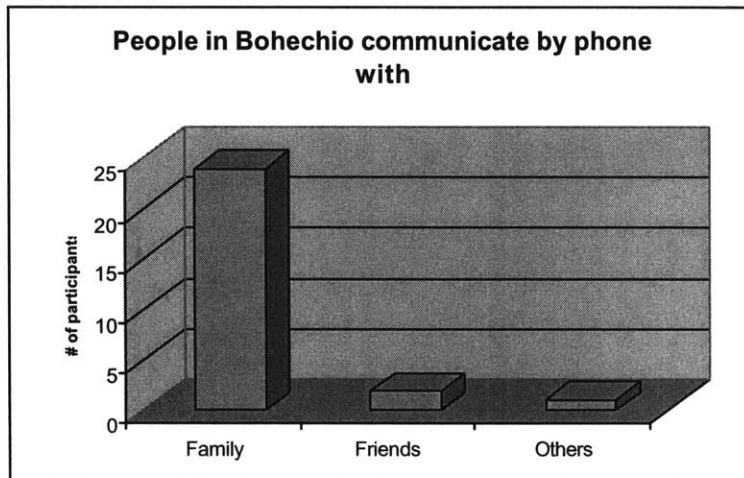
After introducing the participant to Convivo, the interviewer collected basic information about the participant. This included information on age, occupation, and education. The sample characteristics of the twenty-five participants in the Convivo evaluation are as follows:

- Gender: Eleven were men, fourteen were women
- Occupation: Seventeen were students, three were teachers, two were farmers, one was a housewife, one was the LINCOS gatekeeper, one was an electrician
- Age: Seventeen were between thirteen and twenty years old, two between twenty and thirty, four between thirty and forty, and two above forty
- Education: Two began but did not complete primary school, three completed primary school, two completed technical careers, three attended the university, and fifteen are currently studying secondary school.
- Literacy Level: Participants were asked to self-report their reading and writing skills. Only two people who did not complete primary school reported that they could read and write correctly. Although the study is not focused on literacy levels or skills this information was collected to later on determine if some results and user preferences correlate with literacy level.

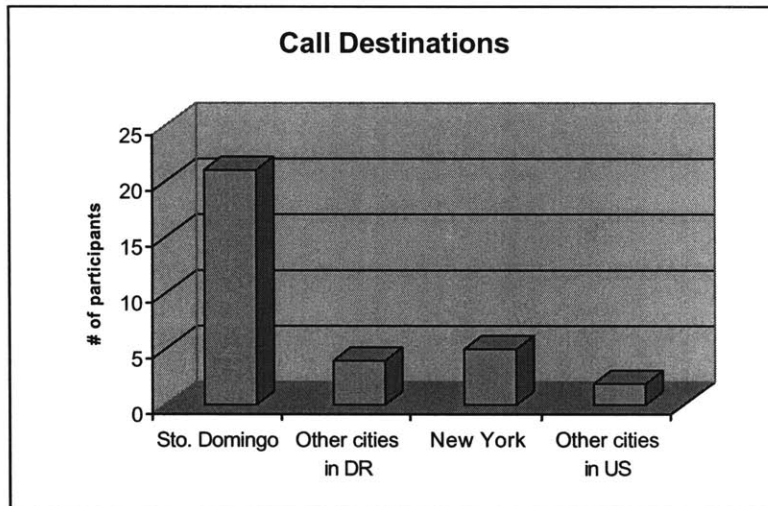
### **5.4.2 Use of Telephone**

Several questions were asked about participant's use of the telephone. Questions asked the participant's monthly expense in phone services, number of calls made/received weekly, and with whom they communicate.

Primarily, the people from Bohechio use the phone to communicate with their relatives who live outside the village (Figure 30). Only three participants mentioned that they use the phone to call people who are not members of their family. The main destination of calls is Santo Domingo, the country's capital, followed by New York (Figure 31).



**Figure 30. People with whom community members communicate by phone**



**Figure 31. Destinations of Telephone Calls**

On average, participants reported spending \$100 Dominican Pesos (approximately US\$5.5 dollars) every month in phone services (Figure 32). Sixteen of the participants consider the telephone service expensive, while the rest consider it as fair.

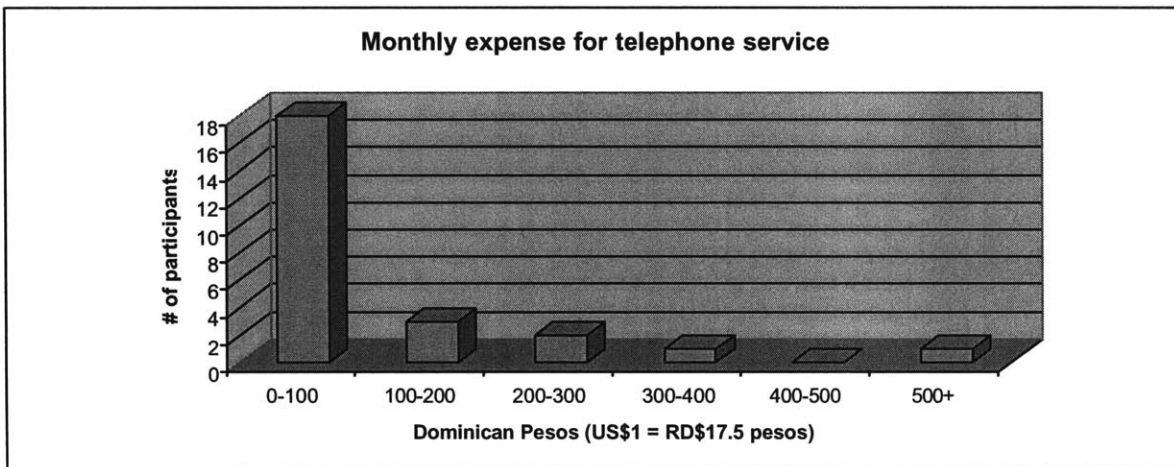
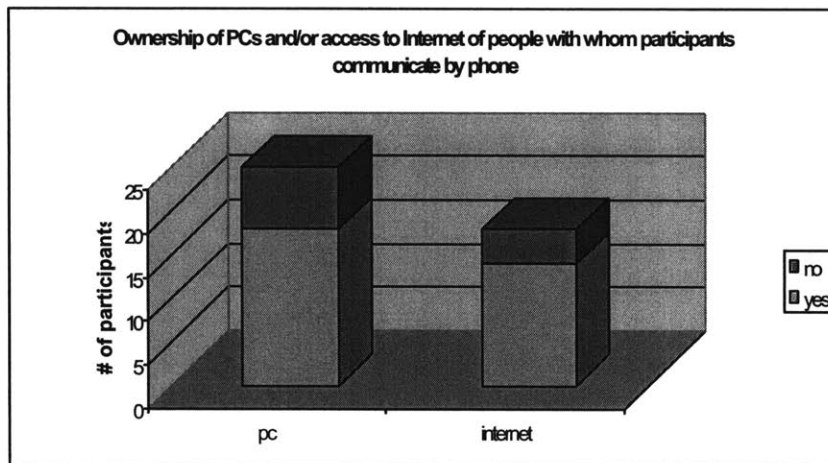


Figure 32. Monthly expenses for telephone service

In the Dominican Republic 80% of Internet users live in the capital, Santo Domingo (UNDP 2001). As previously mentioned, the main destination for incoming or outgoing calls is Santo Domingo. Therefore, there is a high probability that in the future communication links between Bohechio and Santo Domingo could become mainly computer based, given the fact that there are more computers than telephones in Bohechio.

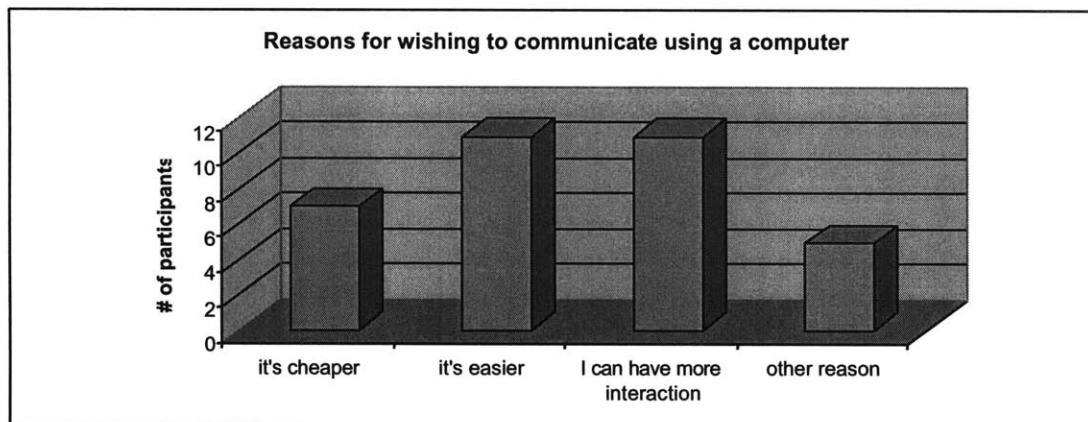
For that reason, in order to explore the potential connection that can be established between participants and the people with whom they communicate, participants were asked to tell if their family in the capital has access to a computer either at home or work. Eighteen of them said that their relatives do, either at work or at home. Out of those, fourteen have Internet access (Figure 33). This data shows the potential that an application like Convivo could have, and how users like the ones in Bohechio would utilize Convivo.





**Figure 33. Ownership of PCs and/or access to Internet of people with whom participants communicate by phone**

Participants expressed a great interest and desire to communicate with their families using computers. According to them, it is cheaper and there is opportunity for more interaction when a computer is used for communication purposes (Figure 34).



**Figure 34. Reasons given for wishing to communicate using a computer**

Also, they think that communicating using a computer gives them the possibility to communicate more often and for longer periods of time. They also mentioned that using the computer for communication purposes it is important and good for the community, and in that way they could learn how to use a computer.

### 5.4.3 Use of Computers

The interviewer also gathered data about the use of computers in Bohechio. To assess participant's computer skills and experience, the interviewer asked if the participant had used computers before. If so, the participant was asked if he had used the computer for communication purposes and the applications he had used to do so. Lastly, the interviewer asked if the participant knows that it is possible to use a computer as a telephone.

Although twenty-one of the participants reported that they had used a computer before, only eleven of them could be considered intermediate or experienced users (Figure 35). This conclusion was made based on the observations that Ezequiel and Adaver gathered while participants were using Convivo during the experiment sessions.

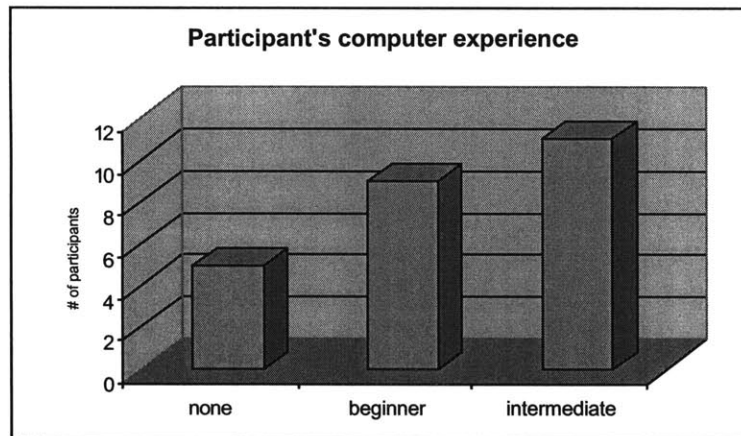


Figure 35. Participants' computer experience

People from Bohechio are already using the computer for communication purposes. Email is the most used application, followed by Microsoft's Instant Messenger (Figure 36). The LINCOS Governmental Office selected MSN Messenger as the official communication application for all the LINCOS telecenters in the Dominican Republic. Therefore, MSN Messenger is installed on every computer in all LINCOS. LINCOS administrators all over the country and the governmental office are in communication by email, and text instant messaging when it is required.

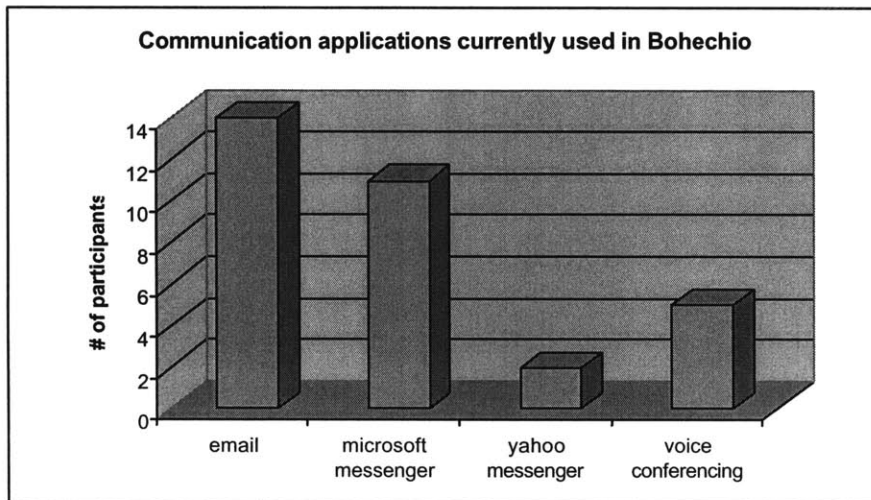


Figure 36. Communication applications currently used in Bohechio

Eighteen subjects claimed to know that it is possible to use a computer in a similar way as a phone. However, only five participants have used a voice over IP application.

## 5.5 Statistical Report and Analysis

Along with providing insight on the background of the participant, all the information previously presented was used to analyze results from the Convivo experiments. Non-parametric statistical analysis was performed to study correlations between participant characteristics and their preferences. In particular, results were analyzed using the chi-square test. Both the chi-squared value ( $\chi^2$ ) and the p-value (p) of the relationship are reported. The chi-squared value calculates the discrepancy between the observed data, which might be related, and a null hypothesis that assumes that there is no relationship in the data. The p-value reports the probability that the correlation calculated by the chi-squared value can be obtained by chance. For example, correlations with a p-value below 1% are considered stronger than those with a p-value of 20%.

Pair wise correlations were calculated between all dependent and independent variables gathered. The dependent variables in the study are: gender, age, education, literacy level, and computer skills. The independent variables are all the opinions gathered during the Convivo experiment

section. Statistically significant correlations were only found between computer skills and independent variables. Those significant results are described in following sections.

## **5.6 Experiencing Convivo**

The Convivo Experiment section was the main section of the user study. The experiment involved the participant in a conversation with one member of the research team by means of Convivo. Both of them had a computer linked to an intermediate station that emulated varying network conditions while they were conversing (see Section 6.4.1). These varying conditions triggered the Interface-Adaptation mechanism producing changes to the graphical user interface. Thus, participants had the opportunity to use all the voice communication modalities available in Convivo. During the experiment, a member of the research team was seated next to the participant to provide assistance as needed.

The Convivo experiment section was conducted in two parts. In the first one, Convivo's Interface-Adaptive mechanism was setup to work in automatic mode. In the second part, it was setup to work in semi-automatic mode. Every section was ten minutes long, so participants were able to experience Convivo for a total of twenty minutes.

### **5.6.1 Network Dynamics Emulated**

Prior to the user study sessions, Bohechio's Internet connection characteristics were obtained from PING tests and measurements taken with the Internet Connection Benchmark 2000 (ICB 2000) network analyzer tool [16]. The PING tests measured the round-trip delay and packet loss from Bohechio to hosts located in other LINCOS sites in the Dominican Republic, Santo Domingo, and the United States. ICB 2000 measured the bandwidth available. The results from these tests are shown in Table 6.

From LINCOS Bohechio to	Round-trip Delay (ms)			Packet Loss %	Bandwidth kbps
	Min	Max	Avg.		
LINCOS Banica, DR	620	1132	650	3-35	24-56
Santo Domingo Government Server	651	1241	669		
MIT website	641	1420	782		
Yahoo.com	560	801	585		

**Table 6. Bohechio's Internet connection characteristics**

The tests conducted showed that packet loss rate is normally in the range of 3-10%, but it can go up to 35%. The bandwidth available for Bohechio according to ICB 2000 measurements is on the range of 24-56 kbps with an average of 36kbps. As mentioned in Chapter 3, there is one satellite Internet connection available in Bohechio which is shared between nineteen computers.

In order to evaluate the Interface-Adaptation mechanism different network conditions were emulated during Convivo experiment sessions. The sequence of network dynamics emulated was selected with the goal to activate the Interface-Adaptation mechanism, which will suggest transitions from one communication modality to another. Therefore, the network performance that participants experienced during the session went from good to poor quality, and vice versa.

The sequence of conditions emulated reproduced network dynamics similar to the ones that a user would experience when using a VoIP application from Bohechio. However, if the entire emulation sequence would be based on Bohechio's Internet connection characteristics (see Table 5) the Interface-Adaptation mechanism would have not suggested the use of the full duplex modality at all. For that reason, unusual changes in network parameters were added to the emulation sequence. In that way, the participant would have the opportunity to experience the three communication modalities and all the possible transitions.

The sequence of network conditions emulated was the same for all participants, it is shown in Figures 37 and 38, delay and packet loss were varying while bandwidth remained constant.

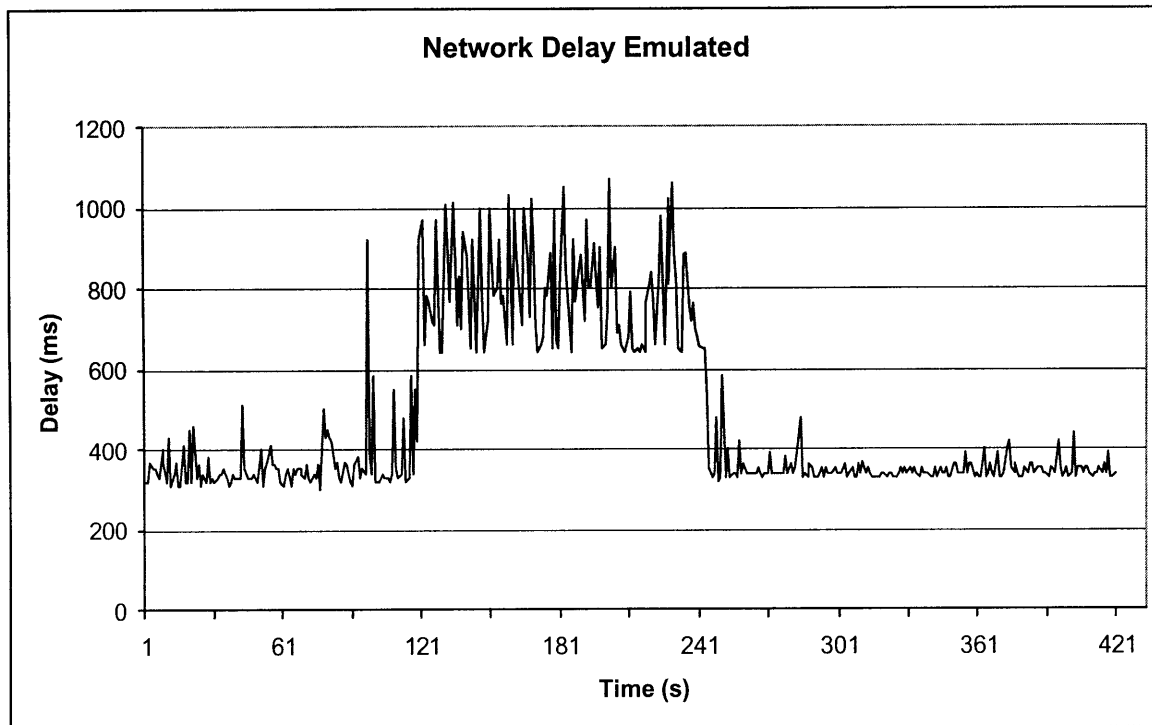


Figure 37. Network delay emulated during Convivo experiment

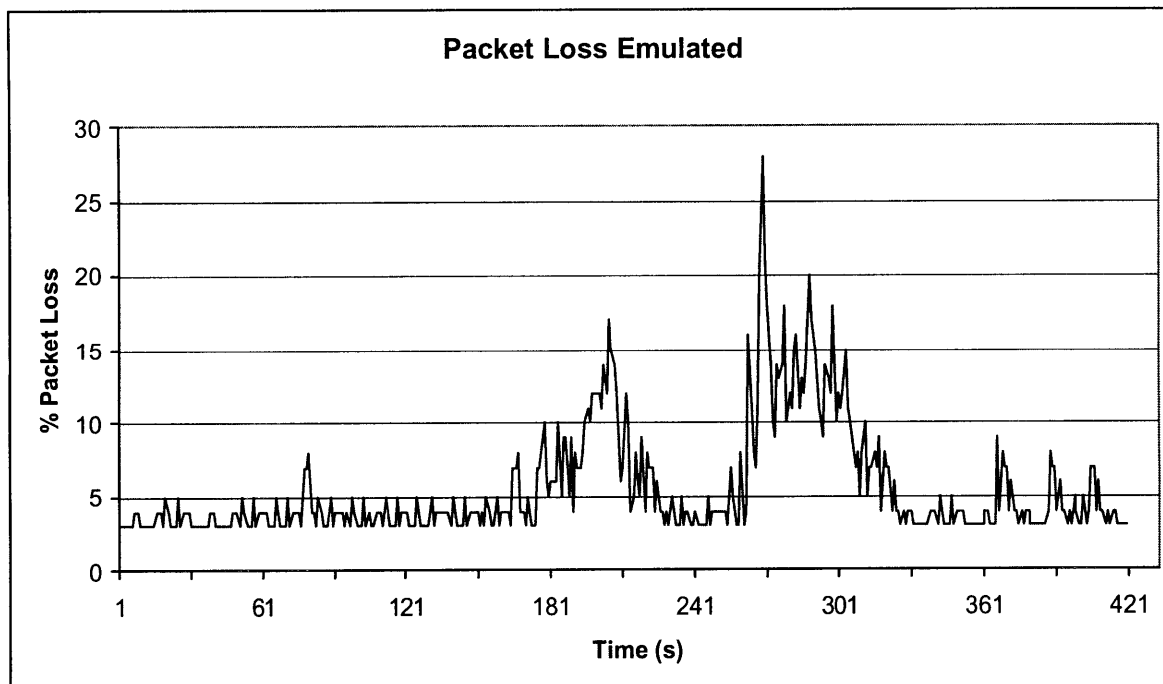


Figure 38. Packet loss emulated during Convivo experiment

### 5.6.2 Participants Perceived Quality

Participants were asked to evaluate the perceived voice and network quality. Participants list the perceived variations (Figure 39). Given the high latency emulated it was expected that several participants would mention had perceived long delay. Surprisingly, only one participant declared that he perceived long delay.

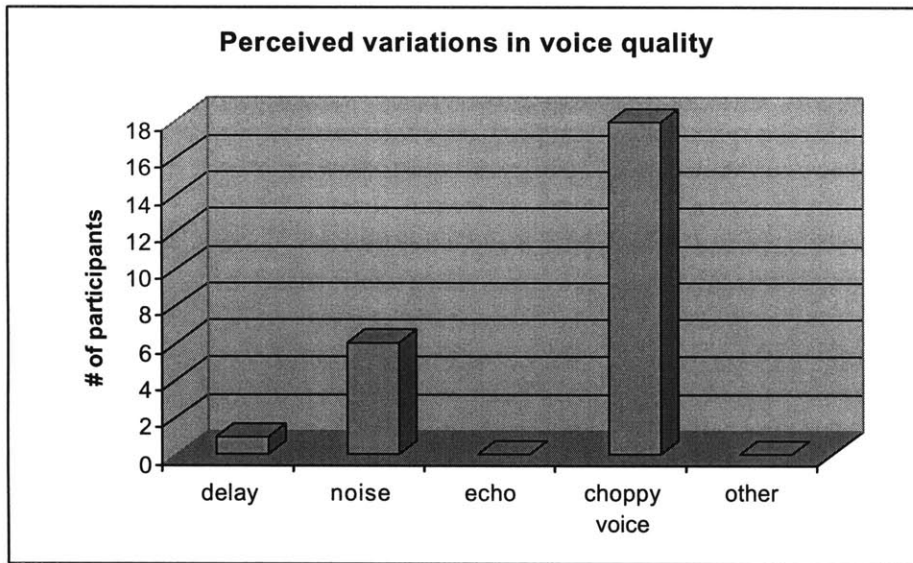


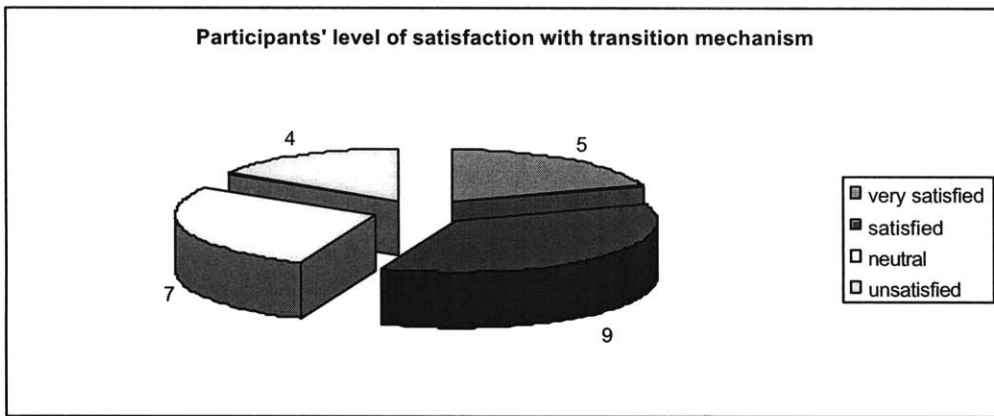
Figure 39. Perceived variations in voice quality during Convivo experiment

### 5.6.3 Interface-Adaptation Mechanism

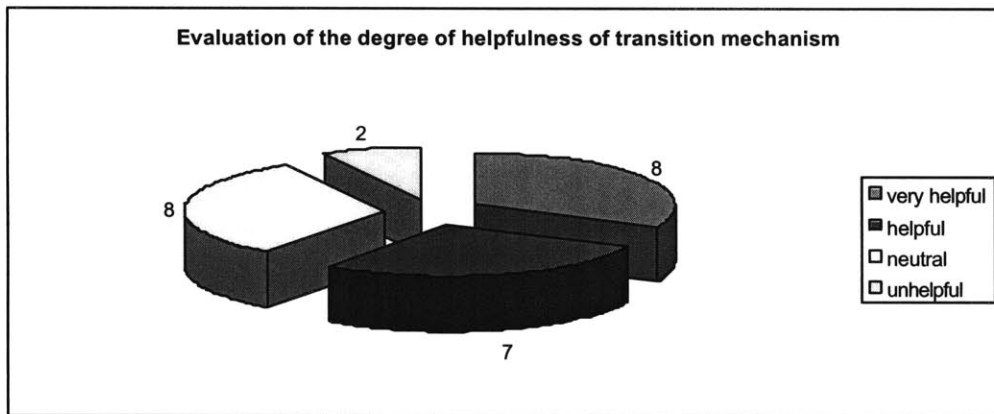
The main goal for the user study was to evaluate the Interface-Adaptation mechanism. Closed questions were used to determine whether or not interface changes were needed and helped participant to communicate as network performance degraded. Opinions of the level of satisfaction with the mechanism are rated as very satisfied/ satisfied/neutral/ unsatisfied. Other questions gather opinions on what actions should the application perform when quality does not improve, and when should the system make a transition. Also, participants were asked to mention the transitions that were the most helpful and unhelpful to maintain communication interaction as network quality changed.

- **Level of Satisfaction**

In general, participants were satisfied with the transition mechanism (Figures 40 and 41). Fourteen participants expressed satisfaction because the system helped them to communicate as the network performance degraded. Eleven of them thought that transitions were necessary due to a poor perceived quality. Only four participants mentioned that they do not want any transition at all. Seven of the participants did not know or did not express any level of satisfaction.



**Figure 40. Participants' level of satisfaction with transition mechanism**



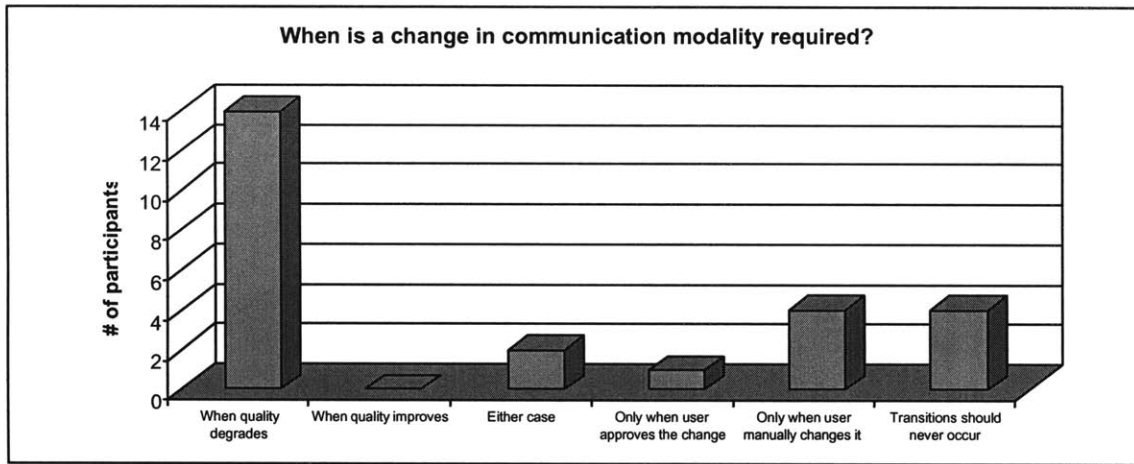
**Figure 41. Evaluation of the degree of helpfulness of transition mechanism**

Although participants were told before the experiment that the user interface might change during the test, facilitators commented that people were surprised when a transition occurred.



- **When to Adapt**

Most participants agreed that Convivo should perform a change on the communication modality when quality varies (Figure 42). This change should happen immediately when network performance varies according to nineteen participants. Only four participants suggested that the Interface-Adaptation mechanism must wait at least 30 seconds before adapting the interface to check if network performance is still poor.



**Figure 42. When is a change in communication modality required?**

Opinions of what should the system do when quality varies show a small correlation with participants computer skills (Figure 43,  $\chi^2 = 14.39$  p = 0.15). Some participants would prefer the system to change immediately and automatically, others would like to be advised prior to a transition, and some others would prefer to be advised and to be able to allow or decline a transition. After observing participants actions during the experiment, it was clear that users with more computer experience and skills would prefer either a semi-automatic transition mechanism or to manually control what communication mode to use.

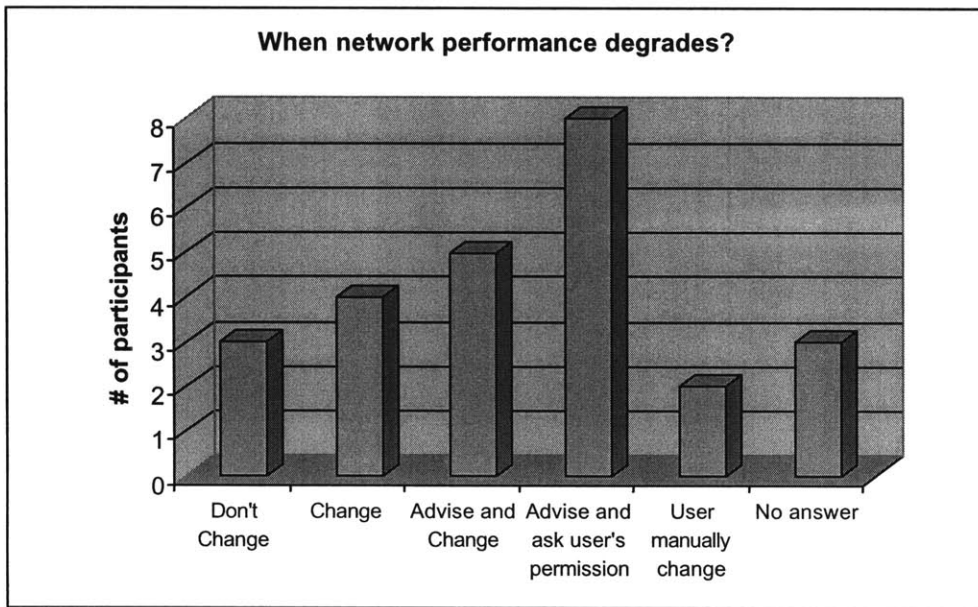


Figure 43. When network performance degrades, what should the application needs to do?

- **Transitions Preference**

The transition from full duplex to voice messaging modality was mentioned by sixteen participants as the change that helped the most as the network quality changed. No other transition was mentioned as helpful, and nine participants did not express any preference.

Eleven participants pointed that transition from full duplex to half duplex did not help at all as the network quality changed. The rest of the participants either expressed that there is not a specific transition they did dislike or did not know.

- **Transition Mechanism Preference**

After experiencing Convivo operating in automatic and semi-automatic transitions, a new set of questions were asked. The questions were oriented to collect participants' preference about the transition mechanism. Participants are asked if the semi-automatic transition mode is better or worse than the automatic mode. The participant is also asked to express which mechanism he would be prefer if the only modes available in Convivo were automatic and semi-automatic.

A total of twelve participants were satisfied with the semi-automatic mode, seven of them thought it was much better than automatic transitions and the other five thought it was better (Figure 44). Only one participant stated that semi-automatic mode was worse than automatic. Fifteen participants preferred the semi-automatic mode to the automatic-mode.

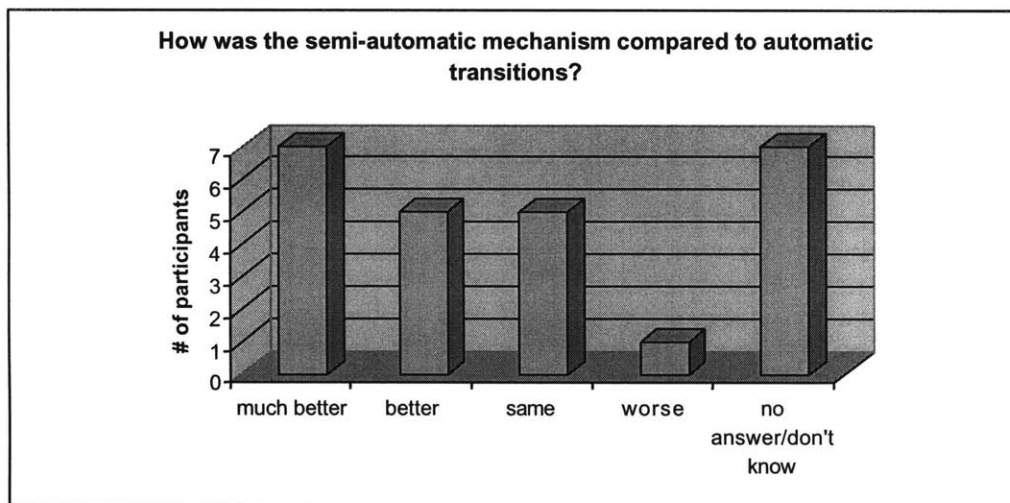


Figure 44. Semi-automatic vs. automatic transitions

Although participants did not experience Convivo with the Interface-Adaptation mechanism in manual mode, they were asked whether or not they would like to manually control transitions. Finally, participants stated what transition would be the best setup for Convivo if the three modes (automatic, semi-automatic, and manual) were available. The final question of the Convivo experiment section asked participants to state what was their favorite voice communication modality.

Nine participants would prefer to manually control the Interface-Adaptation mechanism (Figure 45). The study analyzed data to find correlation between user computer skills and interface-adaptation mechanism preference. A significant correlation was observed between computer skills and transition mechanism mode preference ( $\chi^2 = 13.86$   $p=0.03$ ).

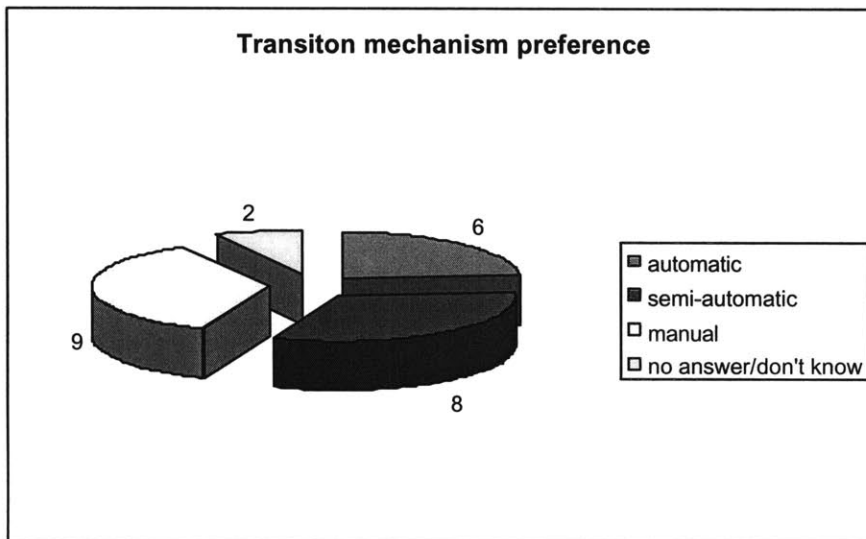


Figure 45. Transition mechanism preference

#### 5.6.4 Communication Modalities

This section presents qualitative assessments and quantitative results on the usability of Convivo interface and communication modalities, as well as results of communication modality preference.

- **Usability**

All the participants described the full-duplex modality as very easy and simple to use because the only action the user has to do is talk. In contrast, several participants had difficulties using the half-duplex modality. According to the facilitators who assisted participants during the experiment, the push-to-talk button from the half duplex modality was hard to use for almost everybody. From their observations we have the following data: thirteen had difficulties all the time, four showed minor difficulties just at the beginning, and only eight users use the button easily.

The main pathologies showed by participants during half-duplex modality were: 1) participants released the button before they started talking; 2) participants thought that after they pressed it

the first time, the button would be activated and would remain active until they finished talking; 3) participants kept the button pressed after they finished talking, they did not released the button. Facilitators' opinion is that most of the problems with the push-to-talk button were because it was the first time that participants were using the application, and in a minor scale, were due to a lack of control over the mouse.

Participants had less problems with the voice messages interface than with the half-duplex. They showed fewer difficulties using the interface. The only pathology detected was that some of them started to use the record button thinking it worked as the push-to-talk button.

- **Usability Test**

The System Usability Scale (John Brooke) was administered to each participant to assess the overall user response to Convivo. Results from the test are summarized below.

- Participants would like to use Convivo frequently
- Participants thought that Convivo is easy to use
- Strong correlation between computer skills and people's thoughts about if they would need to learn other things before they could get going with Convivo ( $\chi^2 = 20.02$   $p = 0.01$ )
- Participants think that they would require support of a person the first time they try Convivo
- Participants agree that most people from Bohechio would learn to use Convivo

While administering the SUS, some participants came up with comments about Convivo usability. They thought that most illiterate and elder people from Bohechio would need to learn how to control the mouse before using the system. Also, some participants thought that some elder people would require the support of a technical person to be able to use Convivo. In general, participants acknowledged that Convivo would be more difficult to use for illiterate and elderly people.

- **Communication Modality Preference**

During both parts of the experiment, all participants had the opportunity to use the three communication modalities. Nineteen participants preferred the full duplex modality as their favorite form of communication. Six participants chose the voice messaging modality as favorite. It is important to note that there was not an open and complete dislike about the half duplex modality (Figure 46).

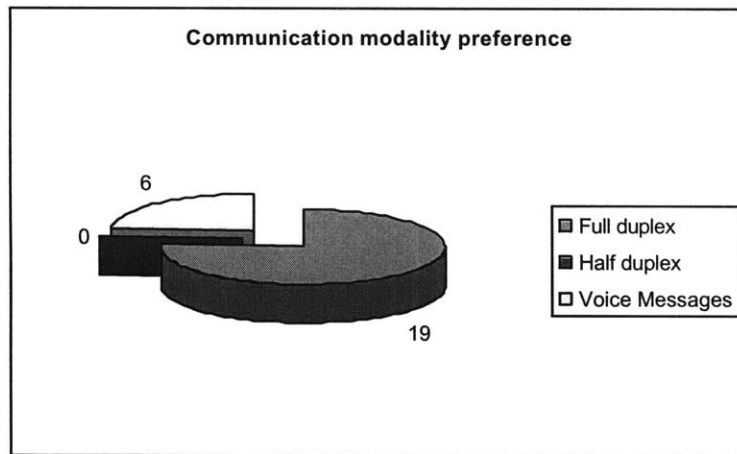


Figure 46. Communication modality preference

The problem with half duplex modality was that participants found the push-to-talk button somehow difficult to use. Therefore, given the option they would prefer to use either full duplex or voice messages. Some participants expressed openly that they did not dislike the half duplex modality. Others think that it requires the user to pay closer attention to the screen. One of the participants who had never used a computer before said that half duplex and voice messages were hard to use but they were very interesting and helpful modalities.

## 5.7 Graphical User Interface Perceptions

Bohechio user study gave the opportunity to get feedback about Convivo's graphical user interface. During the experiment, every participant had the opportunity to give opinions and suggestions about Convivo's GUI. Also, a one-hour participatory design session was conducted

with five of the participants to discuss issues about the GUI. The results of the design session and suggestions from participants during the interview are shown in this section.

### **5.7.1 Conversation Graphical Representation**

Every participant liked and understood completely the graphical representation of conversation provided by Convivo. Participants welcomed with enthusiasm and joy the portrait feature of Convivo. They were extremely excited about the possibility to see a picture of them but more importantly the portrait of the person with whom they were communicating. They expressed that the portrait feature would make the conversation more “familiar”. Participants said it is very important to know more about the person with whom you are talking, the picture is an important step to achieve this goal.

Using the talk-aloud observational evaluation method, each participant had to speak aloud a commentary of what was happening on the user interface. All users could identify themselves correctly on the interface. In half duplex modality, they could describe who was talking and who was listening. In the voice-messaging mode, they completely described who was composing a message, when this was sent, and when the message was received and listen by the other peer.

### **5.7.2 Network Status Indicator Ideas**

Each participant was asked to give his opinion about the network status indicator. Not everybody came up with an idea. Only ten participants elaborated a design for it. The semaphore indicator that would turn the green light on for good quality network and would turn the red light on for poor quality network was mentioned by three participants. Other ideas suggested were:

- Use a text representation: “B” for good, “M” for bad (the B and M are the initial letters for the Spanish words for good and bad respectively)
- Use a symbol: “-“ for good, “X” for bad
- Use an arrow: white arrow for good, red arrow for bad
- Use a prerecorded message to inform the user about network status

- Use a road metaphor: present a nice highway for good network, dirt road for poor network
- Use an image or picture: clear image for good, distorted image for bad

### **5.7.3 Participatory Design Session**

The evaluation process of CKS conducted in Bohechio provided important recommendations for future software developments targeted to rural communities. One of those recommendations was that communities must be part of the software design methodology. Communities must participate in designing and building the technologies.

The evaluation of Convivo in Bohechio took into account that important recommendation. In order to understand the requirements and desires of the users in Bohechio, a collaborative design session was conducted. The session was based on the Focus Group method, which is a method often used by market researchers to assess reaction to products. The focus group consists of a group of volunteer subjects and a group leader, who is member of the evaluation team.

The goal of the focus group session in Bohechio was to collaborate with community members to get opinions and ideas about Convivo's graphical user interface (GUI). Four people took part in this one-hour session. The session was mainly focused on the GUI but it was also open to suggestions of features that users would like to see implemented and issues around the use of Convivo in the context of a community like Bohechio. The results gathered are presented in the following paragraphs.

The group agreed that graphical representation of conversation presented by Convivo is totally understandable. They suggested that man and woman have to be represented differently, there has to be one representation for man and one for woman.

The group questioned what would happen if a user does not have a portrait. The group suggested that Convivo should allow users to select alternative representations if they do not have a portrait. They would prefer to use a toolbar for this purpose.



Another important issue for the group was the visual presence of the application. They mentioned that it was important to design Convivo's GUI with brighter and more attractive colors. "People like brighter colors, if you can do it more people will like to use your application" the group commented.

Related to the transition mechanism, the group recommended the use of audio indications prior to a communication modality change. In addition, they would like to be able to configure the mode in which the Interface-Adaptation mechanism will work: automatic, semiautomatic, or manual.

Two other ideas about network status indicator came up during the participatory design session. The first one was to use a similar representation to the one that cellular phones have for signal strength. The other idea suggested the use of a weather metaphor, use icons that represent a nice sunny day for good network performance and icons that represent a cloudy day or storm for bad performance. Besides these two ideas, the semaphore representation was mentioned again, which seems to be the one that most people from Bohechio would rapidly understand.

Convivo's communication modalities were also part of the discussion. The group acknowledged that for some users it would be difficult to use the push-to-talk button if they do not have any computer skills. Nevertheless, the group thinks that the button is part of that form of communication and that it has to work in that way. Users will get used to it. In order to make that modality better, they think that audio feedback has to be used to indicate a change of turn to talk. No comments were made related to the other two communication modalities.

Unexpectedly, the main topic related to communication modalities was the addition of a fourth modality. The group thought that Convivo might include a text modality as the fourth form of communication available for the user. Two factors were mentioned to support this idea. First, text is an option when network performance is extremely poor because it consumes less network resources. Second, due to the fact that computers in Bohechio are in a common space with other computers sitting closely, some users would like to have more privacy and would not like other

people surrounding them to hear what they are saying. For that reason, text modality is seen as an option if a user wants full privacy in his conversation. However, the group mentioned that the addition of text modality would have the risk of transforming Convivo into a text instant messenger. Convivo could lose its fundamental characteristic, voice communication. At the end, the consensus of the group was that text modality should be included just as a fourth option only available in presence of extreme poor network conditions and only for literate users.

Finally, the group suggested to include multiparty conferencing functionality and stressed the importance of being able to retrieve voice messages that were sent by other users when one is offline.

## **5.8 Directory Service**

After experiment with Convivo, participants were asked to use the directory service to perform two simple queries. The directory service experiment includes the participant using the graphical directory interface. Participants are asked to do two query exercises. The first one requires the participant to find all the males who live in Bohechio and work at LINCOS. The second requires the participant to find all the children who live in Bohechio.

At the end of the two query exercises participants are interviewed. First, they are asked to rank how difficult it was to use the directory. Second, they are asked to list the categories they think would be important for the community and they would like to see on the. Third, the participants are asked to state if they would like to be represented with a self-portrait, and if they have any concerns if the picture is available to every user who access the directory service. Finally, participants are asked to mention any additional comments or suggestions for the directory user interface design.

This section presents qualitative results, observations and suggestions made by participants about the directory service experiment.

It is important to note in first place, that there are no directories available in Bohechio, neither from close communities nor from cities of social and economic relevance for Bohechio, like the towns of San Juan or Santo Domingo. People from Bohechio are not used to search or browse any kind of tool for finding people. As presented early on the chapter, people from Bohechio communicate mainly with their family, and in some cases with friends. Therefore, family and friends were the most demanded categories that should appear on the directory interface.

Participants experienced some difficulty in creating queries that involved more than one field. Participants were able to add at most one field to the query panel. They needed assistance to complete both queries that they were asked to perform as part of the experiment. Double-clicking was the preferred form to add subgroups to the query panel.

Given public access to pictures and basic information is not seen as a problem for participants. All participants welcomed with enthusiasm the idea of including a self-portrait picture on the directory. Moreover, all of them expressed not to have any concern if their picture and basic personal information (name, occupation, age, location) is available for all Convivo users that search the directory.

Several participants saw the directory service user interface not only as a tool to locate people to talk with but also as a more powerful tool. For example, three participants suggested adapting the directory service to provide a service to find jobs in other communities. Two other participants thought that the directory interface is an opportunity to learn more from other regions. They suggested that the directory should present more information about the communities shown. Adaver Anderson, who helped during the interviews, thinks in similar way. The directory service user interface has the potential to be a learning tool. According to Adaver, people can learn about different communities and places all over the world.

## Chapter 6

### Conclusion

This thesis has presented Convivo, a VoIP application designed to provide reliable voice communication for poor quality networks, especially those found in rural areas in the developing world. Convivo introduces an original approach to maintain voice communication interaction in presence of poor network performance: an Interface-Adaptation mechanism that suggests adjusting the application user interface to conform to one of three voice communication modalities. Recognizing that a large number of people in developing countries are illiterate, Convivo presents a user interface to support users with a range of literacy and computer skills. In addition, this thesis proposes a novel graphical user interface for searching and browsing an LDAP directory service, which contains basic information of the individuals contactable through Convivo.

Convivo was evaluated by means of a user study in Bohechio, a rural agricultural community in the Dominican Republic. The study investigated the Interface-Adaptation mechanism, Convivo's user interface, and the addressing scheme proposed. The study found that Interface-Adaptation decisions helped users to maintain voice communication interaction in presence of poor network performance. The transition from full duplex to voice messaging modality was the most helpful change when network quality degraded, while a change from full to half duplex was perceived as unhelpful. The study found that users with more computer skills would like to manually control transitions based on feedback provided by the application and their own perceived voice quality.

Full duplex modality is by far the most liked communication modality because is extremely simple to use and similar to the only form of communication modality that people from Bohechio are used to, the telephone. Nevertheless, few participants mentioned voice messaging as their preferred communication modality. Convivo's user interface was also evaluated. All participants liked and understood completely the graphical representation of conversation that appears on the interface. In particular, the portrait feature was welcomed enthusiastically. Directory service was evaluated as well. The study found that people from Bohechio are not used to consult any kind of directory or phone books. Participants showed some difficulty to create queries that involved more than one field. Family and friends were the most demanded categories to appear on the directory interface.

Convivo provides a voice communication application that is an initial component to develop more elaborated communication systems for rural areas in developing countries. Expanding research and development of appropriate communication applications can help to supply the need for communication, and to reduce the isolation that these communities face.

## **6.1 Future Work**

The following is a list of future works that would contribute to develop a Convivo client with the potential to be fully adopted in rural communities from developing countries:

- Conduct a longer user study that involves people from different communities. The network of LINCOS telecenters in the Dominican Republic is a perfect scenario for such study. A prolonged study would provide valuable feedback about the performance of Interface-Adaptation under real conditions, and about users' transitions and communication modalities preferences. The study would give the opportunity to gather more opinions about the user interface, and features desired by the communities.
- Explore and incorporate other communication modalities to Convivo interface. During the evaluation in Bohechio, participants suggested the incorporation of text-chat as a fourth communication modality. Other modalities that would be interesting to test are videoconferencing, and text-to-speech.

- Propose a login mode that will allow illiterate users to access the system. Smart cards, fingerprint recognition, or iconic passwords are some possible options.
- Develop a web-based version of Convivo. Web browsers are so ubiquitous that are installed almost on every computer in the world. Currently, Convivo is a software application that would require the users to download a setup program to install Convivo. This approach could be complicated for some users. Access to Convivo would be easier if the application runs from a web-page.
- Develop a version of Convivo that runs on handheld devices, such as the Simputer [53] or the Compaq Ipaq [18].
- More research is required to define appropriate icons for the user interface. For example, the network status monitor representation has to be changed for a semaphore icon, according to what the participants of the Bohechio user study suggested.
- Enable a voice messages MailBox for every user. The current implementation of Convivo does not allow users to send voice messages to users that are not online.
- Implement the entire H323 protocol stack to allow Convivo users to place calls to the Public Switch Telephone Network.
- Enable multiparty conferencing.
- Conduct more participatory design sessions to obtain new designs for Convivo's user interface.
- More research is required to determine the organization of the directory service, and the icons to be used to represent the categories of the directory.
- Implement new ways to query the directory service. For example, use speech interfaces to enable voice based queries.

Future work in voice communication applications for rural areas in developing countries should focus on extending VoIP services. Although wireless networks and mobile phones are expanding considerably, access to communication will continue to be one major need, not satisfied in remote areas of these countries. The development of applications should address how to link computer devices in rural areas with fixed and mobile phones in urban areas with systems that are easy-to-use, at low cost, and accessible to everybody. It is important to encourage the creation of new online communication environments designed specifically to increment the

access to specialized information relevant to the community. For example, providing access to educational resources, to technical assistance, and to new economic markets. The development of collaborative communication tools would have a strong impact in fields like agriculture, education, and health.

The main problem that needs to be addressed is to find the applications that would make a real economic and social impact on the communities that would attract all of its members to use information and communication technologies. Closer and prolonged research has to be conducted in rural communities to understand the needs and problems of these communities in order to translate those findings on applications that truly impact and empower people.

If applications are developed to aid communities like Bohechio, the communities must themselves participate in designing and building technologies. Expanding application methodology design and creating tools that allow communities to create their own solutions will accelerate more application development.

# References

- [1] Andren J., Hilding M. and Veitch D., "Understanding End-to-End Internet Traffic Dynamics," submitted to Globecom '98, Sydney, November (1998)
- [2] Best M. L., Burnes L., Escobedo M., Shakeel H. "Village Area Network, Bohechio". To appear in the IFIP WG 9.4 Conference. Bangalore, India. 2002
- [3] Boger Y., "Fine-tuning Voice over Packet services" <http://www.protocols.com/papers/voip2.htm>
- [4] Bolot J.C., "Characterizing End-to-End Packet Delay and Loss in the Internet," J. High Speed Networks, vol. 2, 1993, pp. 305–23.
- [5] Bolot J.C. and Vega-Garcia A., "The Case for FEC-Based Error Control for Packet Audio in the Internet," ACM Multimedia Sys., 1997.
- [6] Borella M. S. et al., "Analysis of End-to-End Internet Packet Loss: Dependence and Asymmetry," Preprint, 1997.
- [7] Cohen, K.M. (1982): Speaker Interaction: Video Teleconferences Versus Face-to-face Meetings. Proceedings of Teleconferencing and Electronic Communications, University of Wisconsin, Madison, pp.189-199
- [8] Compaq, "Handhelds Home," Compaq, 2000. See <http://www5.compaq.com/products/handhelds/>
- [9] Cox R. V., "Three New Speech Coders from the ITU Cover a Range of Applications," IEEE Commun. Mag., Sept. 1997.
- [10] Cruickshank H., Sun Z., Carducci F., Sanchez A., "Analysis of IP voice conferencing over EuroSkyWay satellite system", *IEE Proceedings Communications*, vol. 148, no.4; Aug. 2001; p.202-6
- [11] FAO, "The Internet and Rural Development: Recommendations for Strategy and Activity" September 1996
- [12] Goyal P. et al., Integration of Call Signaling and Resource Management for IP Telephony, IEEE Network, Vol. 13, No. 3, May/June 1999, p.24-32.
- [13] Hagsand, O.; Hanson, K.; Marsh, I., "Measuring Internet telephony quality: where are we today?" *Global Telecommunications Conference GLOBECOM '99*, vol. 3, 1999; p.1838 –1842
- [14] Handlye M: "Session initiation protocol". IETF RFC 2543 (March 1999)
- [15] Hoyler A., "Methods for Evaluating User Interfaces" University of Sussex at Brighton, Cognitive Science Research Paper No. 301
- [16] Internet Connection Benchmark 2000. See <http://www.byzantine.nl/icb>
- [17] IP xStream, Voice and Fax Over IP News and Market Research. See <http://www.iptelephony.org>
- [18] IEEE Communications Magazine, Issues on Internet Telephony, vol. 38, no. 4, Apr. 2000
- [19] IETF RFC 2251, "Lightweight Directory Access Protocol (v3)," December 1997.
- [20] ITU: "Challenges to the Network: Internet for Development", Geneva, 1999
- [21] ITU-T Recommendation G.102: 'Transmission performance objectives and Recommendations', Nov. 1988
- [22] ITU-T Recommendation G.114: 'One-way transmission time', Feb. 1996
- [23] ITU Rec. G.723.1 "Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s" Mar. 1996
- [24] ITU Rec. G.729, "Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)," Mar. 1996.
- [25] ITU Rec. G.729 Annex A, "Reduced Complexity 8 kbit/s CS-ACELP Speech Codec," Nov. 1996.
- [26] ITU Rec. H323, 'Visual telephone systems and equipment for local area networks which provide non-guaranteed quality of service' Geneva, Switzerland, May 1996.
- [27] ITU-T Recommendation P.562: 'Methods for subjective determination of transmission quality', Aug 1996



- [28] Jensen C., Farnham S. D., Drucker S. M., Kollock P., "The effect of communication modality on cooperation in online environments", *Conference on Human Factors in Computing Systems. CHI 2000. The Future is Here, ACM, 2000*; xvii+588 p.470-7
- [29] Kingsley L., "IP Telephony Via Satellite: A Technology Primer" *Internet Telephony Magazine* Dec 2000 <http://www.tmcnet.com/it/oe1200.htm>
- [30] Kostas T., Borella M., Sidhu I., Schuster M., Grabiec J., and Mahler J. Real-time voice over packet switched networks. *IEEE Network*, 12(1):18--27, Jan/Feb 1998.
- [31] Little Intelligent Communities. See <http://www.lincos.net>
- [32] Mukherjee A., "On the Dynamics and Significance of Low Frequency Components of Internet Load," *Internetworking: Research and Experience*, vol. 5, 1994, pp. 163–205.
- [33] Nguyen T., Yegenoglu F., Sciuto A., Subbarayan R., "Voice over IP service and performance in satellite networks", *IEEE Communications Magazine*, vol.39, no.3; March 2001; p.164-71
- [34] NIST Net network emulator. See <http://www.antd.nist.gov/itg/nistnet/>
- [35] O'Conaill, Brid, Steve Whitaker and Sylvia Wilbur (1993): *Conversations over Video Conferences: An evaluation of the Spoken Aspects of Video-Mediated Communication: Human-Computer Interaction*, vol. 8, pp.389-428
- [36] Paxson V., End-to-end Internet Packet Dynamics, *Proc. ACM Sigcomm '97*, (1997)
- [37] Preece, J. (Ed.). (1993). *A Guide To Usability: Human Factors in Computing*. Suffolk, United Kingdom: Addison-Wesley.
- [38] Puetz, J., "Integration of Satellite and Terrestrial." *Satellite Communications*, November 1999, p. 27
- [39] Reynolds R.J.B., Rix A.W., "Quality VoIP – an engineering challenge", *BT Technology Journal*, vol. 19, no.2; April 2001; p.23-32.
- [40] Richardson, D., Paisley, L. *The First Mile of Connectivity, Food and Agriculture Organizations of the United Nations*, 1998
- [41] Richardson, D., Ramirez, R., & Haq, M. *Grameen Telecom's Village Phone Programme in Rural Bangladesh: a Multi-Media Case Study*. Guelph Ontario, Canada: Canadian International Development Agency, 2000
- [42] Richardson D., "The Internet and Rural & Agricultural Development: An integrated approach", *Food and Agriculture Organizations of the United Nations*, 1997
- [43] Rix A.W., Hollier M. P., "The perceptual analysis measurement system for robust end-to-end speech quality assessment", *IEEE International Conference on Acoustics, Speech, and Signal Processing Proceedings, ICASSP '00*, vol. 3, 2000; p.1515 –1518
- [44] Rix, A.W.; Hollier, M.P.; Gray, P., "Predicting speech quality of telecommunications systems in a quality differentiated market" *6th IEE Conference on Telecommunications*, 1998; p.156 –160
- [45] Robust Audio Tool. See <http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/>
- [46] Rose, A., Shneiderman, B., & Plaisant, C. (1995). *An Applied Ethnographic Method for Redesigning User Interfaces*. *Conference Proceedings on Designing Interactive Systems: Processes, Practices, Methods, & Techniques*, 115 – 122.
- [47] Ruhleder K., Jordan B., "Co-constructing non-mutual realities: delay-generated trouble in distributed interaction", *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, vol. 10, no. 1; 2001; p.113-38
- [48] Saban G., Braun R. M., "Wireless local loop and packet radio technology for developing communities", *Proceedings of the 1998 South African Symposium on Communications and Signal Processing, COMSIG '98*. 1998; p.1-6
- [49] Sanghi D. et al., "Experimental Assessment of End-to-End Behavior on Internet," *Proc. IEEE INFOCOM '93*, Mar. 1993, pp. 867–74.
- [50] Schulzrinne H. et al, "RTP: a transport protocol for real-time applications", *IETF RFC 1889* (1996)
- [51] Shakeel, H. "Community Knowledge Sharing: An Internet Application to Support Communications Across Literacy Levels" M.S. Thesis, Massachusetts Institute of Technology (2001)
- [52] Shakeel, H., Escobedo, M., & Garcia-Romero, B. (2000). *LINCOS Bohechio Investigation Results*. Unpublished Manuscript

- [53] Simputer. See <http://www.simputer.org>
- [54] Sun Z., Cruickshank H., "Analysis of Voice Conferencing over Geostationary Satellite Systems", IEE Seminar on Satellite Services and the Internet (Ref. No. 2000/017), 2000; p3/1 – 3/5
- [55] Tang, John C. and Ellen Isaacs (1993): Why do users like video? Computer Supported Cooperative Work: Journal of Collaborative Computing, vol. 1, no. 3, pp. 163-196
- [56] Thorne D., "VoIP – the access dimension", *BT Technology Journal*, vol. 19, no.2; Apr. 2001; p.33-43
- [57] United Nations Development Programme. (2001). Human Development Report 2001. Retrieved from the World Wide Web: <http://www.undp.org/hdro/report.html>
- [58] Watson A., Sasse A., "Measuring perceived quality of speech and video in multimedia conferencing applications" Proceedings of the sixth ACM international conference on Multimedia, Sept. 1998 p55-60
- [59] Weiss, M. and Hwang, J. Internet Telephony or Circuit-Switched Telephony: Which is Cheaper? School of Information Science, University of Pittsburgh, Pittsburgh, PA, Sept 1999; see [www2.sis.pitt.edu/~mweiss/papers/itelv3b.pdf](http://www2.sis.pitt.edu/~mweiss/papers/itelv3b.pdf)
- [60] Yeong W., Howes T., Kille S., "Lightweight Directory Access Protocol". IETF RFC 1777