INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Performance Issues in Next Generation Wireless Multimedia Networks

By

Anthony Malczanek

A dissertation submitted to the Graduate Faculty in Computer Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2004

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 3115271

UMI®

UMI Microform 3115271

Copyright 2004 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This manuscript has been read and accepted for the Graduate Faculty in Computer Science in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Prof. Syed V. Ahamed

12/12/03
Date

Chair of Examining Committee

Prof. Theodore Brown

1/8/04
Date

Executive Officer

Prof. Michael Anshel

Prof. Ali Ghozati

Prof. Michael Kress

Dr. Victor Lawrence

Supervisory Committee

Dr. Shri Bali

Distinguished Guest

THE CITY UNIVERSITY OF NEW YORK

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Abstract

Performance Issues in Next Generation Wireless Multimedia Networks

By

Anthony Malczanek

Adviser: Professor Syed V. Ahamed

Today, the need for increased mobility puts an additional emphasis on wireless communications. Wireless technologies bring several obvious advantages in our everyday activity, like: mobility, scalability, ease of installation, reliability, affordability.

The ubiquitous use of wireless communications will represent an important step ahead toward the final goal of communications: the delivery of any amount of information, anywhere, anytime to anyone is entitled to.

This thesis has two distinctive parts:

- In the first part I make a presentation of the main steps in the evolution of wireless communications. Both wireless telecommunications and data communications are analyzed. A particular emphasis is put on the merge of

the 3 types of wireless communications and the emerging multiservice networks.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- In the second part of the thesis I modeled and simulated different types of wireless multiservice networks, using OPNET simulation software and I explored the performance of these type of networks.

I modeled three types of wireless multriservice networks: ad-hoc, infrastructure, and wireline routed networks. As multimedia communication, digitized and packetized voice is generated.

The study gives answers specific questions for wireless multimedia communications, like:

- as physical layer, is Direct Sequencing or Frequency Hopping more suitable
- as voice codec, is G.711 or G.729 more suitable
- what is the maximum number of wireless users the wireless network can accept without degrading the quality of voice transmission.
- what technologies and in what degree, can solve quality of service in a wireless multimedia network.

The research can be continued with new challenges like:

- measuring the impact of different combinations of data transport over voice transport
- studying the impact of new multimedia applications like digitized and packetized video transmission, over wireless networks

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Content

Introduction		1
Chapter 1.	A short history of wireless Communications	3
Chapter 2.	Wireless transmission	8
	1. Radio Wave propagation	
	2. Modulation	
	3. Access Technologies	
	a1. FDMA	
	a2. Spectral Efficiency of FDMA	
	b1. TDMA	
	b2. Spectral Efficiency of TDMA	
	c1. CDMA	
	c2. Spectral Efficiency of CDMA	
	4. Capacity of Communication Channel	
	5. Impairments in Wireless Transmission	
	6. Architecture of a Cellular Radio System	
Chapter 3.	First Generation Wireless Networks	29
	1. Advanced Mobile Phone System (AMPS)	
	2. Digital AMPS (IS-54B)	
Chapter 4.	Second Generation Wireless Networks	36

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. Time Division Multiple Access (TDMA – IS-136)

2. Global System for mobile Communication (GSM)

3. Code Division Multiple Access (CDMA – IS-95)

4. Internetworking Standard IS-41

5. BS to MSC Communication standard IS-634

Chapter 5. Third Generation Wireless Networks 74

1. Key features

2. Phase 2+G

3. 3G Wireless Systems

4. Universal mobile telecommunication System (UMTS)

5. CDMA2000

Chapter 6. Beyond 3rd Generation Wireless Networks 101

1. Next Generation Wireless Network Architecture and Services

2. Next Generation Wireless Network Regulatory Issues

Chapter 7. Wireless Intelligent Networks 106

1. Intelligent Networks (IN)

2. Advanced Intelligent Networks (AIN)

3. Wireless Intelligent Networks (WIN)

4. Evolution of WIN networks

Chapter 8 Wireless Data Communication Networks 130

1. Classification

2. Wireless ATM

3. IrDA

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4. IEEE 802.11 Standard for Wireless LAN

5. IEEE 802.11 Sub-standards

- IEEE 802.11a
- IEEE 802.11h
- IEEE 802.11b
- IEEE 802.11g
- IEEE 802.11e

6. HyperLAN2

Chapter 9. Mobile IP 169

1. Components of Mobile IP Network
2. Phases of Mobile IP
3. Mobile IP Security Issues

Chapter 10. Multiservice Wireless Networks 179

Chapter 11. Wireless VoIP 189

1. Voice over IP-based Networks (VoIP)
2. QoS Issues for VoIP
3. QoS solutions for VoIP
4. Voice Codecs
5. Wireless VoIP solutions
6. Security solutions for Wireless VoIP networks

Chapter 12. Modeling and Simulation of Wireless Networks 207

1. System Modeling and Simulation
2. Modeling Process

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. Model Engineering

4. Models of Computation

5. Network Simulators

6. Simulation models with OPNET

Chapter 13. Study of Performance Issues in Wireless Multimedia Networks 231

1. Wireless network modeling, using OPNET simulator

2. Elements of the OPNET WLAN simulation model

3. Simulations for the Study of Performance Issues in Wireless VoIP Networks

4. Simulation Scenarios

Conclusions 373

Appendix A. Queuing Theory 376

Bibliography 397

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Figures

Figure 2.1. Architecture of a Cellular System 25

Figure 5.1. 3rd Generation Wireless Network - General Architecture 86

Figure 5.2. CDMA 2000 Architecture 99

Figure 7.1. Intelligent Network General Architecture 110

Figure 7.2. Wireless Intelligent Network (WIN) Basic Architecture 123

Figure 7.3. WIN with a Stand-Alone HLR 124

Figure 7.4. WIN with a merged SCP/HLR 127

Figure 9.1 Mobile IP Components 170

Figure 9.2. Reverse tunneling in Mobile IP 176

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Tables

Table 11.1 Coding Algorithms - main characteristics 199

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Introduction

In the today's world, our activity relies heavily on the capability to share information and resources. That's why we witnessed a longstanding technological revolution in telecommunications. Today, the need for increased mobility puts an additional emphasis on wireless communications. Wireless technologies bring several obvious advantages in our everyday activity, like:

- Mobility: wireless networks allow users real-time access to information from anywhere in the organization, in the country or even in the world, in movement or in a fixed place, without the necessity of a physical connection to the network, thereby dramatically increasing productivity;

- Reliability: the inexistence of wiring creates less maintenance and connection debugging problems for both users and network operators;

- Ease of Installation: Wireless networks do not require expensive and time-consuming cable installations, especially useful in difficult-to-wire or remote areas;

- Affordability: The installation, service and product life costs could be significantly lower than in case of wired networks, especially in environments that require a high mobility;

- Scalability : Wireless communication networks can be more easily configured and re-configured to accommodate a variety of settings and/or number of users.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Classifications

Wireless communications can be divided in different classes/categories, depending on the classification criteria used. We can distinguish the following classification criteria:

- Fix / mobile
- Centralized / decentralized / distributed
- Cellular / Local loop
- Analog / digital
- Narrowband / broadband
- Voice / data / multimedia

The development of wireless networks is subject to several technical limitations like: transfer, capacity, quality, range etc.

The environment where wireless networks can be used are classified in 5 categories:

1. Home-cell environment for in-house applications;
2. In-building systems using picocells
3. Urban areas using microcells
4. Suburban and remote areas using macrocells
5. Global environment, using satellite communications

The ubiquitous use of wireless communications represent an important step ahead toward the final goal of communications: the delivery of any amount of information, anywhere, anytime to anyone is entitled to.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 1. A Short History of Wireless Communications

Wireless communications rely on several fundamental discoveries made in the field of electric waves transmission. The paternity of wireless communications is disputed between the Russian Alexander Popov and the Italian-Irish Gullielmo Marconi.

In May 1895 Popov is accredited with sending and receiving a wireless signal across 600meter distance. In 1897 he demonstrated wireless communications between a land station located in Kronstadt, Russia and a similar station on the "Africa" navy cruiser.

Gullielmo Marconi was 22 years old when in late summer of 1895 made the first demonstration of wireless transmission at his family mansion in Pontecchio, near Bologna, Italy. In December 12 , 1901 he made the first wireless telegraph transmission over the Atlantic Ocean. Wireless communication brought many beneficial contributions in transportation, communications, security and increased mobility. It represents the first important step toward globalization of the world.

A first major application of wireless communications in the US was made in the field of public safety: in 1928, police departments in Detroit, Michigan and Bayonne, New Jersey and the Connecticut State Police were among the first to use in their patrol cars the radio-telephone service, the same one which improved the safety of ocean-going ships.

If these still huge devices were easy to use in ships with no obstacle around, they were too large for street cars running in an uneven landscape, around tall buildings.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

An other important step in radio communications came with the invention, in 1935 by Edwin Howard Armstrong, of the Frequency Modulation (FM). FM brought a tremendous improvement in the quality of the transmission and reduced the bulk of radio equipment.

During World War II the superiority of FM communications became obvious. At the same time, the strategic value of wireless communication on the battlefield stimulated companies like AT&T, Motorola or General Electric, to focus on refining mobile and portable communications. Motorola's FM, first Handie-Talkie and later Walkie-Talkie, is an important example of wireless product developed during the war years and carried over into peacetime use.

Already around 1946 the first commercial wireless phone service became available in Saint Louis, Missouri and AT&T's Mobile Telephone Communication service (MTC) in the New York City. The coming years it became widespread in several big American cities. This system did not use cellular technology, but a single very powerful transmitter, providing coverage 40-50 miles around the base station, and was normally located in the center of a metropolitan area. This system design had great limitations in terms of number of users: just a few radio channels available to carry calls, and limited to only 12 simultaneous callers.

In 1947 at AT&T's Bell Laboratories was invented the cellular technology, in an effort to use the airwaves more efficiently. The limited number of radio frequencies available for mobile service was simultaneously used in multiple low-power transmitters laid throughout a metropolitan area, and "handing off" calls from transmitter to transmitter as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

customers moved around in their vehicles. When more capacity was needed, the area served by each transmitter could be divided again. The technology was implemented by AT&T under the name of MTS (Mobile Telephone Services). This event represented the generalization of wireless technology.

In 1964 AT&T launched an improved version of mobile services, named IMTS (Improved Mobile Telephone Service); this improved version provided full-duplex operation, automatic dialing and automatic channel searching.

Soon after, in a joint AT&T and Motorola effort a new system named AMPS (Advanced Mobile Phone System) was in place, but not yet publicly available.

Always, it took two more decades until the complex call handoff technology was developed, in order to make wireless service ubiquitous, and at the same time for regulators to give approval:

- in 1973, Motorola introduced its revolutionary new DynaTAC mobile phone, a conveniently sized radiotelephone set.

- in 1977, the FCC authorized two experimental licenses - to AT&T in Chicago, and to Motorola and American Radio Telephone Service, Inc. in the Baltimore/Washington, D.C. corridor.

Cellular technology became of public use, first time in 1979 in Japan, with an implementation known today as the Japan System.

North-European countries launched in public use in 1981 the NMT (Nordic Mobile Telephone) system, followed by TACS in Great Britain, in 1983.

As an example of the dimension of use, in 1976 in New York City there were 545 mobile phone users and 3500 in a waiting list.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In the United States, only in 1983 the AMPS system was publicly introduced, first in Chicago, IL. Concomitantly, the FCC announced that lotteries would be used to award licenses in all markets below the top 30 systems. By 1984, Washington, D.C. for example, had two competing wireless providers. By 1990, construction permits had been issued for at least one system in every market in the United States.

At the end of 1990 in the US, there were over 5 million wireless subscribers. A landmark of 10 million subscribers was registered on November 23, 1992.

In August 1993, the FCC received authorization to auction PCS (personal communications service) licensees. The FCC defined PCS as a "broad family" of mobile radio services (as for example as advanced voice paging, two-way acknowledgment paging and data services), which could be used to provide service to individuals and businesses, and would be integrated with a variety of competing networks.

In October of 1993, the FCC concluded that a combination of large Major Trading Areas (MTAs) overlapping smaller Basic Trading Areas (BTAs) would "promote the rapid deployment and ubiquitous coverage of PCS and a variety of services and providers." Three more auctions followed, as the FCC offered 1,972 additional BTA-based licenses. The wireless competition that began in 1983 with two licensees per market was further expanded in 1995 to provide for up to nine carriers per market. The wireless industry has become a model of market competition.

As mentioned in the CTIA website dedicated to the history of wireless communications: "As a result, the wireless industry is also at the forefront as government policy evolves from a substitute for competition to an enabler of competition. Congress adopted the wireless model of competition in lieu of government intervention for the entire

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

telecommunications industry in 1996, recognizing that competition has been the driving force behind the innovation and growth that has characterized wireless telecommunications".

In 2000 in the US there were 86 million wireless subscribers

These days, the rate of subscribers is growing at a rate of 67,082 new wireless subscribers every day, which means about one subscriber every two seconds.

Bibliography

1.1. "Guglielmo Marconi", http://www.wsone.com/fecha/marconi.htm

1.2. "History of Wireless", http://www.wow-com.com/consumer/faq/articles.cfm?ID=101

1.3. http://www.kdischool.ac.kr/library/data/t00054.PDF

1.4. "Russia's Popov: Did he invent radio?", http://www.wsone.com/fecha/popov.htm

1.5. The official U.S. Government publication History of Communications - Electronics in the United States Navy (Bureau of Ships and Office of Naval History, 1963)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 2. Wireless Transmission

Radio waves with different frequencies have very dissimilar propagation characteristics. Consequently, certain parts of the radio spectrum are far more suitable for land-mobile communications than other.

2.1. Radio-Wave Propagation

The part of the electromagnetic spectrum, representing Radio Waves, ranges from 30KHz to 300GHz, corresponding to wave-length from 10km to 1mm.

In free space, radiowaves propagate in straight line and are reflected off objects. Therefore in space the transmitter and the receiver should be in Line-of-Sight (LOS) or the radio wave should be reflected off a conducting object.

Radio waves on the earth are affected by the topology of the ground, atmosphere, natural and artificial obstacles on the ground.

Depending on the frequency of the transmission there are three main propagation means:

- Ground waves: is the part of the radiation which travels in contact with the earth surface being scattered by the topology of the terrain and objects on the terrain.
- Tropospheric waves: is the part of the radiation that is kept close to the earth surface as a result of refraction from the lower atmosphere (troposphere). Tropospheric refraction makes sometimes VHF to travel much longer distances than ordinary ground waves. Tropospheric communication increases with the increase of frequency.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Always at UHF level tropospheric waves cause interference and is not a dependable form of communication.

- Ionospheric (sky) waves: is the part of the radiation that leave the transmitter at angles above the horizontal. Under certain conditions these waves can reach and be reflected by the upper layer of the atmosphere, named ionosphere, and reach the ground again at distances ranging up to 4000km from the transmitter. By successive reflection between the ionosphere and the earth surface, these waves can travel at very long distances, even around the globe. The best suitable waves for this type of transmission are those below 30MHz. It is dependent on the sunspot cycle.

2.2. Modulation

There are three basic ways of modulating a carrier: frequency, amplitude and phase, as well as a combination if these modulation techniques.

The output voltage of a carrier has the following sinusoidal form:

$V(t)=A\cos(\omega t+\theta)$, where:

A – is the amplitude of the signal

ω - is the frequency of the signal, and

θ - is the phase of the signal.

Any of these 3 parameters may be varied in accordance with the information-carrying signal.

2.3. Access Technologies

Traditionally there are two categories of access technologies:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- narrow-band and channelized
- wideband and non-channelized

In a channelized system the total spectrum is divided in a several narrow-band channels. Each radio channel is defined by a pair of channels, for duplex communication.

The channel used to communicate from the base station to the Mobile Station is named forward channel and the channel used to communicate from the Mobile Station to the Base Station is named Reverse Channel. A user is assigned both Forward and Reverse Channel for the duration of a call. Usually the Forward and Reverse Channels are separated by a wide frequency range, in order to keep interference between transmission and reception to a minimum.

For call set-up and control, dedicated Control Channels are allocated.

The most widely used Multiple Access Technologies are:

a. **Frequency Division Multiple Access** (FDMA): The total frequency spectrum is divided in channels. Guard-bands are maintained between adjacent channels to minimize cross-talk between channels. Advantages of FDMA are:

- a capacity increase can be obtained by reducing the information bit-rate and/or using efficient digital encoding
- implementation is technologically simple

Disadvantages are:

- improvement in capacity depends on reduced transmission ratio
- the cost of narrowband filters is higher (because they are not realized, using VLSI)
- maximum bit-rate is fixed and small, implying low flexibility and data transmission power

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- for duplex operation, the MS antenna is provided with a duplexer, limiting size, weight and cost.

Spectral Efficiency of FDMA

In FDMA the multiple access, for speech efficiency, is reduced by the existence of guard-bands between channels and of the signaling channels.

Multiple access is defined as the ratio of total time/frequency domain allocated to voice transmission and the total time/frequency domain available to the system.

For FDMA the multiple access spectral efficiency is:

$$\eta_a = \frac{B_c N_T}{B_w}$$

Where:

η_a = is the multiple access spectral efficiency

N_T = is the total number of voice channels in the area

B_c = is the channel spacing

B_w = is the bandwith of the system

b. Time Division Multiple Access (TDMA)

Data from each user is conveyed in time intervals, called slots. Several slots make up a frame. Each slot is made up of preamble and information bits. The function of the preamble is to provide identification and synchronization. Guard times are used between each user's transmissions to minimize cross-talk between channels.

The advantages of TDMA are:

- allows a flexible bit rate for both high speed and low speed transmission

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- does not require costly narrowband filters
- Flexibility for frame-by-frame monitoring of signal-strenght/bit error rates for hand-off
- Utilizes bandwith more efficiently than FDMA for lack of frequency guard-band
- Each signal is transmitted with enough guard time to accommodate:
  - Time inaccuracies due to clock instability
  - Delay spread
  - Transmission time delay because of propagation distance
  - "Tails" of signal pulses because of transient responses

The disadvantages of TDMA include:

- For handsets, for transmission on the up-link channel, high peak-power is necessary
- Requires a large amount of signal processing for matched filtering, correlation detection, and synchronization with the time-slot.

Spectral Efficiency of TDMA

In TDMA the multiple access for speech efficiency is reduced by the existence of guard-time and synchronization sequence.

Multiple access is defined as the ratio of total time/frequency domain allocated to voice transmission and the total time/frequency domain available to the system.

For TDMA the multiple access spectral efficiency is:

$$\eta_a = \frac{\tau M_t}{T_f}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Where:

η_a = is the multiple access spectral efficiency

τ = is the duration of the time slot

T_f = frame duration

M_t = number of time slots per frame

c. Code Division Multiple Access

In CDMA each user is assigned a unique code sequence (named spreading code) used to encode the information bearing signal in such a way that the entire system bandwith is made available to each user.

Therefore, the used bandwith is much larger than the necessary bandwith required to transmit the information. Such systems are named Spread Spectrum (SS) systems.

A SS modulation technique must fulfill two criteria:

- the transmission bandwith must be much larger than the information bandwith
- the resulting radio-frequency bandwith is determined by a function which is independent of the information signal.

There are three fundamental types of SS systems:

- Direct Sequence Spread Spectrum (DSSS): the information signal is directly multiplied with a high chip-rate spreading code
- Frequency Hop Spread Spectrum (FHSS): the carrier frequency of the information signal is rapidly changed according to the spreading code. FH could be:
  - fast: if the hopping rate equals or exceeds the information symbol rate
  - slow: if two or more symbols are transmitted between frequency hops.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Time Hopping Spread Spectrum: (THSS): the information signal is transmitted in short bursts, where the times of the bursts are decided by the spreading code.
- Hybrid CDMA systems are those systems that employ a combination of two or more of the above mentioned modulation techniques or with a other type of multiple access technique.

We can have the following hybrid systems: DS/FH, DS/TH, FH/TH, DS/FH/TH or CDMA/TDMA. The idea of the hybrid systems is to combine the advantages of different modulation techniques.

The main advantages of CDMA are:

- the ability to tolerate a fair amount of interference
- frequency band assignment is greatly simplified (adjacent cells share the same frequencies)
- flexibility in system design and deployment are significantly improved
- can accommodate more mobile users (than FDMA or TDMA) due to voice activity patterns during two-way conversation
- much higher security level of the transmission

The main disadvantages of CDMA are:

- requires very low synchronization error between the received code sequence and locally generated code sequence, which is very difficult to obtain;
- requires a much more precise power control, which is very difficult to obtain because of feedback delays, imperfect power estimates, errors in the feedback channel and traffic conditions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- if synchronization is poor then self-jamming happens; the despreading operation does not recover the original narrowband signal and also there will be partial-correlation with the codes of other users.

Spectral Efficiency of CDMA

For CDMA the multiple access spectral efficiency is expressed by the formula:

$$\eta = \frac{N_u R}{B_w}$$

Where:

N_u = is the capacity of the system

R = is the information bit-rate (including overhead)

B_w = is the bandwith of the system.

And, the capacity of the system is expressed by the formula:

$$Nu = \frac{\eta_f \eta_b c_d G}{v_f} \times \frac{B_w}{R \times (E_b/I_o)}$$

where:

η_f = is the frequency reuse efficiency

η_b = is the bandwith efficiency factor (< 1)

c_d = is the capacity degradation factor

v_f = is the voice activity factor

B_w = is the bandwith of the system (one way)

R = is the information bit rate (including overhead)

E_b = is the energy per bit of the desired signal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

E_b/I_o = is the bit energy to interference ratio

G = is the number of sectors in cell cite antenna.

2.4. Capacity of a Communications Channel

Let it be a communication channel transmitting S watts of power. The majority of the transmitted power is in a bandwith, noted B_w. We also assume that the only impairment on the channel is thermal noise. The bandwith of this noise is much larger than of the signal bandwith.. If we assume that the distribution of the noise is Gaussian, then the channel is named Additive White Gaussian Noise Channel (AWGN). C.E. Shannon proved that there is a maximum rate at which the information can be transmitted over the channel error-free.. This transmission rate is named the "error-free capacity" of the communication channel. The error-free capacity of the channel is given by:

$$C/B_w = \log_2 [\ 1 + S / (\ N_0.B_w)\ ] = \log_2 [\ 1 + (E_b/\ N_0).(R / B_w)]$$

Where:

C = channel capacity (in bits/sec)

B_w = one-way transmission bandwith (Hz)

E_b = energy per bit of the received signal (Joule/bit)

R = information rate (bits/sec)

$S = E_b.R$ = signal power

N_0 = single sided noise power spectral density (W/Hz).

An ideal communication channel is the one in which information is transmitted at the maximum rate: $R = C$. In this case the formula become:

$$C/B_w = \log_2 [\ 1 + E_b/N_0) \cdot (C/B_w)]$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Solving the equation for E_b/N_0, we get:

$$E_b/N_0 = (2^{\alpha} - 1) / \alpha,$$

Where: $\alpha = C/B_w$.

As a result, for any given rate-to-bandwith, there is a signal-to-noise ratio (SNR) above which, error-free transmission is possible.

Practical communication channels are compared to this ideal channel by choosing a probability error (as for example 10^{-6}) and finding the signal-to noise ratio necessary to achieve it.

There are four conventional ways the capacity of a cellular system can be enhanced:

- adding channels
- sectoring cells
- borrowing channels
- splitting cells.

2.5. Impairments in Wireless Transmission

Due to different physical conditions wireless transmission can be impaired by factors as: noise, multipath fading, Doppler effect, channel interference.

Further, we shortly describe the main possible impairments:

1. Noise: is present in any communication system. The noise is a random process function, usually having a Gaussian probability density. This additive noise factor will cause the amplitude and phase of the transmitted electromagnetic signal to vary.

Classification of noise:

- Thermal noise: is introduced by temperature dependent resistive elements in circuits

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Human made noise: include such sources as automobile ignition, electric motors, mechanical vibrations, industrial spurious emissions etc.

- Path loss: in a free space communication system the path loss is proportional to the square of the distance. More precisely, in decibels, the path loss is given by the equation:

Path loss = -32.45 + 20 * Log(Frequency) + 20 * Log(Distance),

where Frequency is in MHz and Distance is in km.

2. Fading: is a fluctuation in amplitude of the wave cause by multipath transmission. There are two types of fading:

- Flat fading happens when the received signal spectrum is close to the transmitted one except for a fluctuation in amplitude.

Flat fading occurs when the Root Mean Square (RMS) delay spread of the channel is much smaller than the symbol period of the transmitted signal.

- Selective fading happens when the RMS delay spread of the channel is more than about 10% of the symbol period. In frequency selective fading, the channel is dispersive, and the received waveform has intersymbol interference (ISI). In the frequency domain, the channel response can no longer be considered flat, because its amplitude has significant variation and its phase is not linear with frequency. Therefore, the received symbols can no longer be identified individually. In this case, a channel equalization technique is used, which removes the intersymbol interference

Technologies to overcome or reduce impairment

The errors induced by transmission impairment could be overcome or controlled. The main error control technologies are:

1. error prevention

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. error detection
3. error correction

Multipath Characteristics of Radio Wave

The energy of radio wave signals exists in the form of electrical and magnetic fields, varying sinusoidal with time. The two types of fields always coexist and there is a continuous flow of energy from one field to the other.

As the wave arrives to the receiver on multiple path due to reflections and refractions, there is a substantial amplitude and phase fluctuation at the received signal and the signal is subject to fading. Also, in case of mobile receiver, there is a Doppler shift in the received signal.

There are two types of multipath fading:

- The rapid fluctuations caused by the local multipath are known as fast fading , or Raleigh fading. Fast fading is observed over distances of about half the wavelenght.
- The long-term variation in the mean level is known as slow fading, or log-normal fading. Slow fading occurs by movement over long distances, large enough to produce big variations in the overall path between sender and receiver.

Consequently, the received signal s(t) can be expressed as the product of two components: the signal subject to long-term fading p(t) and the signal subject to short term fading r(t):

$$S(t)=p(t)r(t).$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

If the signal is vertically polarized (the electrical field factor is aligned along the z axis) then the field components are: two magnetic field components H_x and H_y and the electrical field vector E_z. These components, at the receiving point are expressed using Clarke's model:

$$H_x = -\frac{E_0}{\eta} \sum_{i=1}^{N} \sin \alpha_i e^{j\varphi_i}$$

$$H_y = -\frac{E_0}{\eta} \sum_{i=1}^{N} \cos \alpha_i e^{j\varphi_I}$$

$$E_x = E_0 \sum_{i=1}^{N} e^{j\varphi_i}$$

where:

E_0 – Amplitude of the N plane wave,

φ_i - Phase angle relative to the carrier phase,

η – Intrinsic wave impedance

A mobile station, moving in a certain direction with a constant velocity v, has the received carrier Doppler shifted, according to the formula:

$$F_d = f_m \cos \varphi = (v/\lambda) \cdot \cos \varphi = v_{eff}/\lambda = (v_{eff} f_c)/c$$

where:

$f_m = v/\lambda$ = maximum value of Doppler frequency

v_{eff} – effective velocity of the vehicle

f_c – the carrier frequency

φ – the path angle with x axis

c – speed of light ($3 \cdot 10^8$ m/s)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The waves arriving from ahead of the mobile station have a positive Doppler shift, whereas those arriving from behind, have a negative Doppler shift.

Delay Spread

Because of multipath reflections, the radio signal follows different paths. Each path has a different path length, which makes that each path to arrive to the destination at a different time. The effect which either gathers or spreads out the signal is named delay spread. In a digital system, the delay spread causes inter-symbol interference (ISI), therefore limiting the maximum symbol rate.

Coherence Bandwith

Coherence bandwith is the bandwith for which either the amplitudes or the phases of two received signals, present a high degree of similarity. Two signals outside the coherence bandwith will fade independently.

The coherence bandwith for two fading amplitudes of two received signals is:

$$\Delta f > B_c = \frac{1}{2 \pi \tau_d}$$

The coherence bandwith for two random phases of two received signals is:

$$\Delta f > B_c = \frac{1}{4 \pi \tau_d}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where τ_d is the average duration of the fading effect.

Propagation Losses in built-up areas

As the mathematical modeling of radio wave propagation in real world is extremely complex, several empirical models to predict propagation losses have been developed. Built up areas are categorized in:

- urban areas, dominated by tall buildings or other commercial or industrial structures
- suburban areas, with residential houses and/or parks or other open spaces
- rural areas, with open farmland, some farm buildings, wooded areas and forests, terrain could be flat hilly or mountainous

As the natural or man-made objects influence radio wave propagation, they were categorized by:

- building characteristics, as density, height, location, size
- vegetarian density
- terrain variation.

In order to build-up simulation models, all these characteristics have to be quantified. The following quantitative measures are defined:

- Building Size Distribution (BSD): as a PDF where the standard deviation indicates homogeneity; small value indicates areas with buildings of uniform size, and high value indicates diversity of range.
- Building Area Index (BAI): defined as the percentage of a 500x500 m^2 test square, covered by buildings, regardless of their heights.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Building Height Distribution (BHD): a PDF of the heights of the buildings in the considered area
- Building Location Distribution (BLD): is a PDF (probability Density function) describing the location of the buildings, in the considered area
- Vegetation Index (VI): is the percentage of the area covered by high vegetation (like trees)
- Terrain Variation Index (TVI).

Several of radio wave propagation models have been proposed. None of them is universally available, but is dependent on an area with certain characteristics, as well as the radio communication technology used.

Most widely known are: Okumura Model, Sakagmi and Kuboi Model, Hata Model, Ibrahim and Parsons Model, W.C.Y. Lee Model.

2.6. Architecture of a Cellular Radio System

Components

The main components of a cellular radio system are:

- **a. Mobile Station (MS):** contains the mobile transceiver, that allows tuning to a particular frequency, designated for its use, by the network.

- **b. Base station (BS):** also known as land station or more precisely Base Transceiver Station, manages the air interface between MS and BS, provides power control operations on the air interface (radio resource management (RR))

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- **c. Base Station Controller:** controls the Base Stations, by off-loading functions from the MSC; it controls the power signals and frequency administration between the MS and BS, and also is responsible for hand-over operations.

- **d. Mobile Switching Center:** is responsible for switching the calls to the cells, interfacing with telephone networks, monitoring traffic, providing accounting, performing testing and diagnostic services, providing back-up and overall network management

- **e. Databases**

  - **e1. Home Location Register:** is a database associated with the cellular operator, storing information about each subscriber, as: accounting information, subscribed services etc.) and also keeping track of a subscriber's location.

  - **e2. Visitor Location Register**: is a database keeping track of the visiting subscribers; the VLR informs the visiting subscriber's HLR when this one logs into the system, and in turn HLR will update its records with the whereabouts of the subscriber, and will send to the visited VLR information about the subscriber.

Note: The overall operation of HLR and VLR is named Mobility Management (MM), because they are responsible for keeping track of the location of each Mobile Station (MS).

  - **e3. Authentication register:** is a database, containing authentication and encryption information on each subscriber; it interacts with HLR and VLR in order to provide these services.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- e4. **Equipment Identity Register (EIR):** is a database containing information on the MS equipment (i.e. manufacturer ID, conformance testing information, and the identity of the assembling factory).

Note: The actual deployment of these four databases is implementation specific.

The figure below represent a typical architecture o a cellular system:

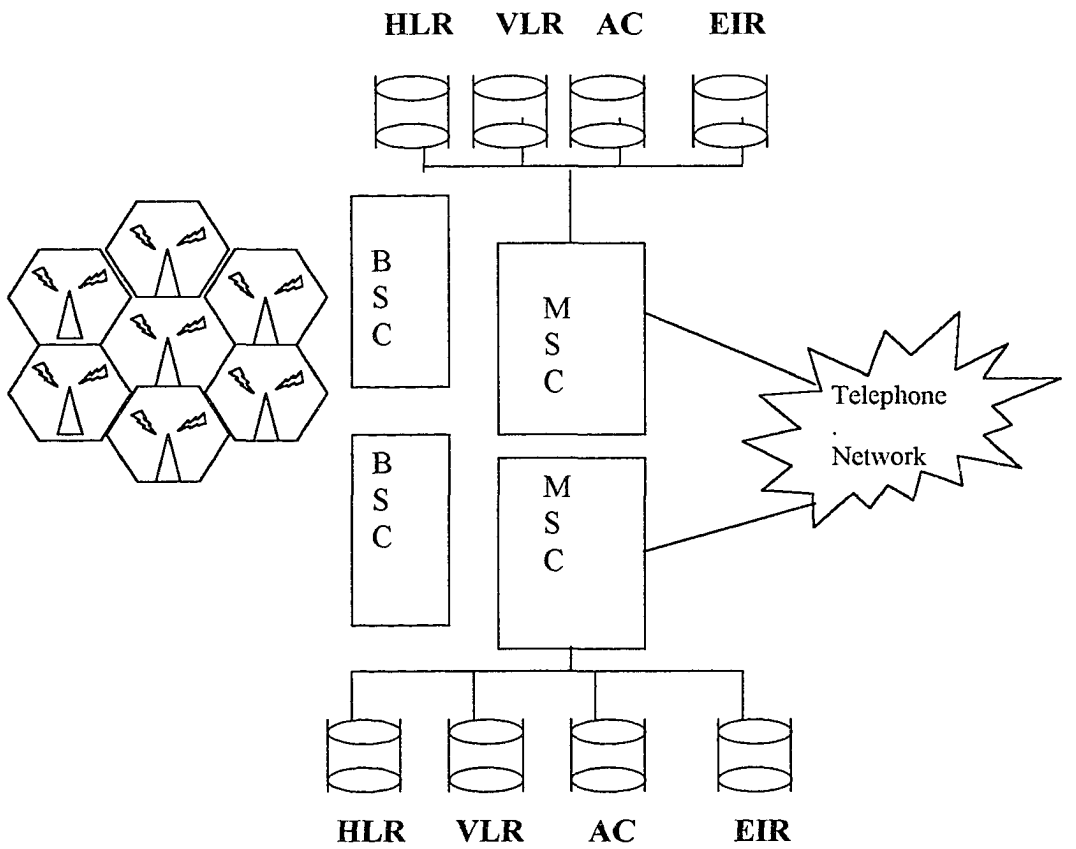


Figure 2.1. Architecture of a Cellular System

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Connection Management Operations (CM)

CM operations are used to set-up and clear-down a call between an MS and another party.

The following sequence of events describe a connection management operation:

1. The calling party accesses the telephone network
2. The call is routed to a gateway MSC
3. The MSC interrogates the HLR for further routing
4. In turn, HLR is eventually interrogating the VLR that is currently serving the user, for routing information
5. VLR is returning routing information to the HLR, which passes it back to the MSC
6. Based on this information MSC routes the call to the terminating MSC
7. The terminating MSC then queries the VLR to match the incoming call with the identity of the called subscriber
8. The terminating MSC sends to the BS a paging request
9. The page is relayed to the called subscriber and the call set-up is completed.

Mobility Management Operations:

1. Upon crossing a cell boundary the mobile station (MS) sends a location update request, containing its identification, to the base station (BS).
2. The message is routed to the new VLR, which sends a query to the MS HLR.
3. HLR stores the MS new location and downloads at the new VLR the subscription information and authentication parameters of the subscriber.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4. The new VLR sends the acknowledgement of the location update, through MSCs, to the originating mobile user
5. The HLR sends a cancellation message to the old VLR, in order to clear the subscriber's data from its database.

Bibliography

2.1. Ahamed, Syed V., Victor B. Lawrence. "Intelligent Broadband Multimedia Network". Kluwer Academic Publishers. Boston, USA. 1997

2.2. Ahamed, Syed V., Victor B. Lawrence. "Design and Engineering of Intelligent Communications Systems". Kluwer Academic Publishers. Boston USA. 1997.

2.3. Balston, D., R. Macario. "Cellular Radio Systems". Artech House. Norwood, MA. 1994

2.4. Clarke, M. "Networks and telecommunications: Design and Operation". John Willey, New York, NY. 1997.

2.5. Garg, Vijay K., Joseph E. Wilkes. "Wireless and Personal Communications Systems". Prentice Hall. New York. 1996.

2.6. "Introduction to Wireless Systems". Wiley and Sons. 1999.

2.7. Paetsch, Michael. "Mobile Communications in the US and Europe: Regulation, Technology, and Markets". Artech House, Boston, MA. 1993

2.8. Pahlavan, K. A. Levesque. "Wireless Information Networks". John Willey. New York, NY. 1995.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.9. Rappaport, Theodore S. "Wireless Communications: Principles and Practice. Second Edition. Prentice Hall. 2001.

2.10. Stallings, William. "Data and Computer Communications". Prentice Hall. 2001

2.11. Webb, William. "The Complete Wireless Communications Professional". Artech House Publishers, Boston USA. 1999.

2.12. Young, Harry E. "Wireless Basics". Second Edition. Intertec. 1996.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 3. First Generation Wireless Networks

3.1. Advanced Mobile Phone System (AMPS)

Advanced Mobile Phone System (AMPS) was standardized in 1976 as a compromise between different inventions made at AT&T and Motorola and conceived to overcome some deficiencies of older wireless systems. It became of public use in 1983.

AMPS is a first generation system used mainly in the US. It is designed to support voice communications, by modulating the 3KHz voice signal onto 30KHz channels in the 800MHz spectrum, using analog signaling technique. on the air interface between MS and BS.

Frequency Allocation

A total of 416 channels are allocated for traffic and control channels. Operator A has allocated 312 voice channels (1-312) and 21 control channels (313-333). Operator B has allocated also 312 voice channels (355-666) and 21 control channels (334-354). A group of 16 voice channels is controlled by one control channel.

- Identifiers:

Three identification numbers are used in the AMPS technology:

a. Mobile Station's Electronic Serial Number (ESN)

b. Mobile Operator's System Identification Number (SIN)

c. Mobile Station's Mobile Identification Number (MIN)

- Control Channels:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

21 channels are available for each operator for control purposes. In idling mode a MS must monitor designated control channels. Some control information are sent over the voice channels, in certain conditions.

- The Forward Control Channel (FOCC) is a continuous stream of data sent from the BS to MS. Three data streams are Time Division Multiplexed (TDM): stream A and B are identified with the least significant bit of the MIN, and the busy-idle stream indicates the current status of the Reverse Control Channel, which can be used by more than one MS.

- On the Reverse Control Channel (RECC), control information is sent from the MS to the BS.

Both channels work at 10kbs.

- Voice Channels

Voice conversation takes place over the Forward Voice Channels (FVC) from BS to MS and Reverse voice Channels (RVC) from MS to BS. The channels are selected by the network and assigned to a MS during a call set-up.

- Using the FOCC, BS informs MS which RVC to use for the call. MS tunes to this channel.

- BS will also switch to the selected FVC for the call.

- Control signaling

Supervisory Audio Tone (SAT) is used to determine link continuity between the MS and BS. BS adds this tone on each FVC prior to the modulation of the signal. Similarly, MS sends the same tone over the RVC to the BS. One SAT signal is sent at least every 250ms. 3 SAT signal are defined: 5970Hz, 6000Hz, 6030Hz. MS is informed about

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

which SAT signal to expect by a special message sent over FOCC, called the SAT Color code (SCC).

- Signaling Tone (ST) is used to perform 4 operations:
  - The "Request to Send" (RTS) signal is sent when the Send button is pushed, and has 400 ms of duration. It is used to allow more data to be input from the keypad.
  - "Alert Signal" is a signal which is continuously sent over the RVC after the MS station was alerted and until the called party answered.
  - Disconnect Signal is sent over the RVC channel by the MS to indicate the termination of the call. The duration of the signal is 1.8 seconds.
  - Handoff Confirmation Signal is sent by the MS, after it receives a message from the network for a handoff operation. The duration of the signal is 50ms.

Communication Procedure

Next, we describe the main steps in the communication procedure.

- Initializations Procedures:

1. MS is activated (power-up) and Receive System Parameters: in which MS receives and configures parameters to use cellular provider A or B

2. MS scans all 21 control channels to receive overhead messages; a channel is selected within 3 seconds. Also during this state MS enters Enbled Status – preferred system is A, or Disable Status – preferred system is B. MS receives the Sid of the serving operator.

3. MS receives a system parameter message over the control channel

4. MS compares its own SID to the SID in the message, identifying the serving operator. If the two SID are not equal parameters are set for roaming operation.

5. MS sends over RECC its MIN, ESN and Sid, in order to identify itself to the system.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6. These parameters are examined by MSC and its VLR to determine the roaming status of the MS

7. BS sends a final control message to MS to verify the initialization parameters

8. MS enters in idle state: MS changes with BS at least every 46.3ms synchronization signals, and keeps informed the network about its mobility status.

- Mobile station Originates a call

The steps, in MS originating a call, are:

1. Through reverse control channel, MS sends a origination message to BS; the message contains Ms's MIN, ESN and phone number and also the order code and order qualification code

2. BS gets the origination message and passes it to MSC for further processing, via IS-41

3. BS sends a control message back to MS, containing the channel number which will be used for voice, as well as the order and order qualification code fields. The message also contains the Set Color Code (SCC) which designates the SAT frequency that is to be used.

4. Now MS switch to voice channel

5. Further, BS sends a forward voice channel (FVC) control message with the SAT signal

6. MS confirms the channel continuity through a reverse voice channel signal (RVC).

7. MSC through BS will confirm MS about the completion of the call

8. Conversation starts on the designated voice channels.

- Mobile Station receives a Call

The steps, in MS receiving a call, are:

1. The identification of the MS is passed to BS through IS-41.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. The information from IS-41 is mapped into a page control message, and additional information, like MIN, SCC, ORDER, ORDQ are added. Also the voice channel number is added, instructing the MS which traffic channel to use for connection

3. MS answers to this message, by returning its MIN, ESN, ORDER, ORDQ fields in a page response message

4. When receiving this message, BS builds a control message containing the SCC field, which informs MS about the frequency to be used on the identified channel.

5. Also BS switches to voice channel

6. At the same time Ms switches to voice channel

7. BS sends an SAT

8. MS answer to this message by a confirmation message

9. Traffic channel is opened and conversation starts on the designated channel.

- Network-side Operations

On the network side, the messages exchanged between BS, MSC and PSTN are a combination of IS-41 and SS-7 messages. The exact sequence in which the events occur depend on the type of the call as well as the concrete implementation of the network.

- Handoff Operations

A handoff operation occurs when the MS moves to another cell. The following sequence of events constitute a handoff:

1. BS-1 handles a call for a given MS

2. BS-1 notices the MS transmission power is decreasing

3. BS-1 sends a handoff measurement request message to its MSC

4. MSC requests a number of BSs to report receiving signal from the MS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.MSC decides which BS is the best choice to take over the call (BS-2) and allocates a traffic channel to BS-2; BS-2 acknowledges the allocation

6. MSC sends a handoff order message to BS-1, which sends a similar signal to MS

7. After confirming, MS receives BS-2 SAT and regenerates back to BS-2 the received SAT, as a continuity check

8. BS-2 confirms the handoff to the MSC

Today AMPS looks like a primitive analog technology, which does not use voice digitization or encryption, but at the time it was conceived, it was advanced and it served the industry well for many years.

3.2. Digital AMPS (IS-54-B)

Digital AMPS is a hybrid air interface, using both first and second generation technology. It was introduced in North America in the early 90-ies to overcome some limitations of the original AMPS technology. It has a dual mode operation in which analog FDMA and digital TDMA are combined.

IS-54-B defines:

- the use of the analog traffic channels
- the use of the digital traffic channels, and
- analog control channels.

Channel bandwith as well as frequency allocation are not changed from the AMPS specification.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

D-AMPS dual mode operation was attractive because allowed operators to migrate to digital technologies on an evolutionary basis. As an intermediary technology, it found its way not only in the US, but in several countries around the world.

Bibliography

3.1. Black, Uyless. "Second Generation Mobile and Wireless Networks". Prentice Hall, Upper Saddle River, NJ. 1999

3.2. Paetsch, Michael. "Mobile Communications in the US and Europe: Regulation, Technology, and Markets". Artech House, Boston, MA. 1993

3.3 Papadimitriou, G. "Wireless Networks". Wiley and Sons. December 2002.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 4. Second Generation Wireless Networks

4.1. Time Division Multiple Access (TDMA), in standard IS-136

IS-136 is a second generation TDMA based technology, built on the D-AMPS operations. It is widely used in North and South America and is of limited use in the Far East. In many aspects is very close to the other second generation TDMA specification, born in Europe, the GSM standard.

The Air Interface

The frame structure consists of 6 slots per frame. Each frame duration is 40ms, and the transfer rate is 25 frames per second. The number of data bits in a slot is variable:

- from network to MS, the length of the data field is 260 bits
- from MS to network, the length of the data field is either 244 or 200 bits, depending on the format of the slot.

IS-136 Identification Numbers

a. MS Electronic Serial Number (ESN): is a 32 bit number, uniquely identifying a MS; it is kept in ROM memory, as any attempt to change it makes the MS unusable

b. System Identification Number (SID): is a 15 bit binary number, uniquely identifying a cellular system. MS is using it during transmission, as the BS can determine under which system are they communicating.

c. MS Mobile Identification Number (MIN): is a 34 bit binary number, representing the 10 digit telephone number of the MS.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Frequency Allocation

TDMA, as specified by IS-136, works both at 800MHz and 1900MHz spectrum.:

- In the 800MHz spectrum the channel allocations are identical to AMPS specifications
- In the 1900MHz spectrum (known also as PCS-1900) the channel spacing is also 30KHz for MS-BS communications, and the transmitter center frequency is as follows:
  - MS to BS: .03n+1850.01
  - BS to MS: .03n+1929.99

    where: n- represents the number of the channel.

Channel Organization

TDMA time-slots are also called channels. Digital channels are organized as follows:

1. Digital Traffic Channels (DTC): are organized into:
   - 1.1 Forward Digital traffic channels (FDTC): are down-link channels
     - 1.1.1 Fast Associated Control Channel(FACCH)
     - 1.1.2 Slow Associated Control Channel (SACCH)
     - 1.1.3 Traffic channels
   - 1.2 Reverse Digital traffic channels (FDTC):
     - 1.2.1 Fast Associated Control Channel
     - 1.2.2 Slow Associated Control Channel (SACCH)
     - 1.2.3 Traffic channels
2. Digital Control Channels (DCCH): are organized into:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.1 Reverse Digital Control Channel (RDCCH): is a special slot, used to send traffic to the BS

2.1.1 Random Access Channel (RACH): is a uplink channel used to gain access to the system, to respond to pages, and send registration messages

2.2 Forward Digital Control Channel (FDDCH)

2.2.1 Short Message Service, Point-to-point, shared, ACKnowledged Channel (SPACH): organized into:

2.2.1.1 Paging Channel (PCH): is used to send various control messages to the MS

2.2.1.2 Access Response Channel (ARCH): is used by the MS to receive various control information as, response to an MS access operation, channel assignment, handoff

2.2.1.3 Short Message System Channel (SMSCH): is used to send short messages to the MS

2.2.2 Broadcast Control channel (BCCH): is organized into:

2.2.2.1 Fast Broadcast Control Channel (F-BCCH): carries general information as: type and number of configured slots, authentication, power level

2.2.2.2 Extended Broadcast Control Channel (E-BCCH): also carries general information as: information on neighbor cells, services supported by neighbor cells

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.2.2.3 SMS Broadcast Control Channel (S-BCCH): is used to deliver short messages.

2.2.3 Shared Channel Feedback (SCF): is used to coordinate random access operations: it informs MS about the outcome of access attempts and which slot to be used for that.

2.2.4 Reserved Channels

Control Channel Selection

The MS is assigned to a certain DCCH through a hashing function. This function hashes the MS to a DCCH based on the number of DCCHs in a cell and the number of slots allocated to a DCCH; this identifier is named PAID. The hashing process consists of obtaining the list of available DCCH in the cell and using the PAID to select the appropriate DCCH and slot, from the list.

Error Control

IS-136 states three error control operations:

- Convolutional coding
- Interleaves speech frames
- Cyclic Redundancy Check

Despite the many conceptual similarities between IS-136 and GSM, the two standards are incompatible and it requires a dual-mode handset for a user to be able to use both systems.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.2. Global System for Mobile communications (GSM)

GSM was initiated by the European Commission in June 1987. The aim was to create a European-wide communications structure.

- GSM-900 is the original implemented system from mid-1980ies.

- DCS-1800 is a updated version of GSM-900. In 1997 DCS-1800 was renamed GSK-1800.

PCS-1900 is also based on GSM-900, and is the North-American version.

GSM Interfaces

GSM standard has four interfaces:

1. Um interface: is the air interface between MS and BTS
2. A interface: is between MSC and BSS
3. A-bis interface: defines operations between BSS and BTS
4. Mobile Application Part : defines the operations between MSC and the telephony network, as well as between MSC and HLR, VLR and/or EIR. It is an extension of SS7 signaling , operating above the SS7 protocol stack.

Air Interface

The Um air interface uses a combination of FDMA and TDMA techniques

The initial GSM-900 operated in the 890-915MHz range for MS to BS transmission and 935-960MHz range allocated for BS-to-MS transmission.

The GSM-1800 similarly uses the 1710-1785Mhz and 1805-1880MHz ranges, therefore increasing the bandwith capacity of GSM-1800 over GSM-900.

Each bandwith is divided in 124 channels with 200KHz spacing between channels (to prevent interference), and as there are separate uplink and downlink channels, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

communication is full-duplex. These FDMA channels use TDMA slots, with 8 slots allocated to a frame. Each slot contains 156.25 bits, with a time slot of .577ms.

Of these bits, 114 are payload and the remaining are used for synchronization and control functions. As there are 8 time-slots, each one takes-up a channel after 4.615

GSM Layers

1. Um interface uses:
   - Q.931+
   - LAPDm
   - Radio Interface
2. A-bis interface:
   - Q.931+
   - LAPD
   - T1/E1
3. A interface:
   - DTAP, BSSMAP
   - SCCP
   - MTP3
   - MTP2
   - MTP1
4. MAP:
   - MAPn
   - TCAP
   - SCCP

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Or : Q.931+, ISUP, TUP

and

- MTP3
- MTP2
- MTP1

Mobile Application Part

As different procedures are used between different components of the GSM network, each-one is identified with MAP-x, where x- is the corresponding procedure.

The table below makes a correspondence between "x" and the corresponding interface:

MAP-B	MSC-VLR		
MAP-C	MSC-HLR		
MSC-D	HLR-VLR		
MSC-E	MSC-MSC		
MSC-F	MSC-EIR		
MSC-G	VLR-VLR		
MSC-H	AC-HLR	(1)	Is different for PCS-1900
MSC-I	UAP-HLR	(2)	Is not defined in PCS-1900

MAP-B is executed between MSC and VLR, when a foreign MSC needs to interrogate a non-attached VLR. (If VLR is integrated into MSC, this protocol is not used).

MAP-C operates between MSC and HLR. Its execution takes place when another MSC needs to query a remote HLR, providing subscriber information to that part of the mobile

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

network through which the home subscriber is roaming: call origination and termination privileges, accounting information, authentication and encryption keys.

MAP-D provides the interface between HLR and VLR, when a VLR/HLR needs to communicate with a non-attached VLR/HLR. It is used to query a HLR by a VLR for a user profile, as well as some security information about the subscriber.

MAP-E is one of the most extensive MAP protocols. It governs the operations between MSCs. It is responsible for the overall control and management of the mobile network, in intersystem Ms roaming and for the SMS service.

MAP-F operates between MSC and EIC in order to obtain the International Mobile Station Equipment Identity (IMEI) value.

MAP-G operates between VLRs. It is used to support roaming and handoff information, about a newly arriving MS.

Functional Planes

Functional planes are used to describe the detailed operation of the GSM across the interfaces.

The following functional planes are described in the standard from top-down:

1. Operations, Administration and Maintenance (OAM): deals with monitoring activities, configuration operation etc.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. Communication Management (CM): is responsible for set-up and tear-down of calls
3. Mobility Management (MM): registration of the mobile station, verification of user with EIR, which services can be used, location management (with HLR and VLR).
4. Radio Resource management (RR): deals with establishing and releasing connections between the MS and the network, involving operations such as: assign and release of channels, manage handover, set cipher mode operation, receive status report from MS, change frequencies,
5. Physical: radio signals carrying the control signals and the voice.

GSM Addresses and Identifiers

GSM use several addresses and identifiers at the air and network interfaces.

- Addressing numbers:

1. International Mobile Subscriber Identifier (IMSI): is used as a unique identifier in a network and is made up of:

- Mobile Country Code (MCC) 3 digits, administered by ITU-T
- Mobile Network Code (MNC): 2 digits, is a unique identifier of a network provider
- Mobile Subscriber Identifier Code (MSIC): 10 digits, represent the costumer identification number.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In order to provide subscriber identity confidentiality instead of the IMSI, a Temporary Mobile Subscriber Identity (TMSI) is used. The number, provided by the VLR after authentication, is valid until the MS is in the same cell.

IMSI is stored in SIM, HLR, VLR and AUC, when a mobile subscriber registers with a certain service provider.

SIM is located in MS, and could be either permanently inserted or written on a removable smart card. In the latter case, the user can use any Mobile Station, allowing a so-called "SIM roaming".

Beside IMSI, SIM contains also:

- phone number,
- Personal Identification Number (PIN)
- Security/authentication parameters.

2. Mobile Station ISDN Number (MSISDN): is a number used to reach a called party in a GSM network. It uses the same format as the ISDN address, based on ITU-T recommendation E.164.

In GSM a call does not identify the called party but rather the HLR. The HLR uses this number to provide routing instructions to other components of the network, in order to reach the called subscriber.

3. Location Area Identifier (LAI): is similar in content to IMSI, but it is used to identify a cell or a group of cells. It is made up of:

- Mobile Country Code (MCC): 3 digits
- Mobile Network Code (MNC): 1 or 2 digits
- Location Area Code (LAC): 10 digits

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.International Mobile Station Equipment Identity (IMEI): is assigned to each Mobil Station (MS) by the manufacturer. It is made up of:

- Type Approval Code (TAC): 3 digits, is given when the mobile unit passes quality control tests
- Final Assembly Code (FAC): 1-2 digits, is given to identify the manufacturing assembly line
- Serial Number (SNR): 9 digits, is a unique number assigned to MS by the manufacturer
- An additional spare digit is available for future assignment.

GSM Channel Organization

- GSM-900 uses two 25MHz bands for the uplink and respectively the downlink. The uplink and the downlink is separated by a 45MHz spacing. Both uplink and downlink are divided in 124 physical channels of 200KHz each, numbered 1-124.
- GSM-1800 uses two 75MHz bands for the uplink and the downlink. The uplink and the downlink are separated by 95MHz spacing. Both uplink and downlink are divided in 374 channels of 200KHz each, numbered between 512 to 885, .
- GSM-1900 uses two 60MHz bands for the uplink and the downlink, separated by 80MHz spacing. Both uplink and downlink are divided in 299 physical channels of 200KHz each, numbered 512-810.

The Absolute Radio Frequency Channel Number (ARFCN) are calculated as follows:

Specific. Freq. Assignment ARFCN Mapping

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

	Up-link	Downlink	Uplink	Downlink
GSM-900	890-915	935-960	890+.2*n	935+.2*n
GSM-1800	1710-1785	1805-1880	1710+.2*(n-511)	1805+.2(n-511)
GSM-1900	1850-1910	1930-1990	1850+.2(n-512)	1930+.2(n-512)

The physical channels are managed through various TDMA techniques, to create logical channels. Logical channels are divided in:

- Traffic Channels (TCH)
- Control Channels (CCH)

The logical channels are organized as follows:

1.Traffic Channels (TCH):

- Full rate
- Half-rate

2. Control Channels (CCH):

- Broadcast Control Channels
- BCCH
- Frequency Correction Channels (FCCH)
- Synchronization Channels (SCH)
- Common control Channels:
- Random Access Channel (RACH)
- Paging Channel (PCH)
- Access Grant Channel (AGCH)
- Dedicated Control Channel:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Stand-alone Dedicated Control Channel (SDCCH)
- Slow Associated Control Channel (SACCH)
- Associated Control Channels:
- Fast Associated Control Channel (FACCH).

Framing Structure

GSM has a hierarchical framing structure:

- a slot has a duration of 0.577ms
- 8 slots are combined in a frame, with a time period of 8*.577=4.615ms.
- 26 frames or 56 frames are multiplexed in a multiframe. The 26 frames multiframe therefore, has a duration of 120ms.
- Superframe is made-up of 26*51=1326 frames, which represent the smallest multiple of all cycles. It has a duration of 6.12 seconds.
- Hyperframe is made up from 2048 superframes, or 2715647 frames. It has a duration of 3hours, 28 minutes, 53 seconds, 760 milliseconds. Within a hyperframe, frames are numbered from 1 to 2715647.

To keep count of the received/sent frames three counters are used:

- T1 – superframe counter, ranges between 0-2047
- T2 – Speech frame counter: ranges between 0-25
- T3 – Signaling frame counter: ranges between 0-50.

With the three counter, the position of each slot on the air interface is uniquely known.

Authentication and Encryption

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

GSM provides authentication by comparing the information stored on the SIM card and from AUC. The authentication is based on A3 algorithm and a private key, noted Ki (stored both on SIM and AUC).

The A3 algorithm from AUC uses the Ki private key and a 128bit random number (RAND) to produce the Signed Response (SRES). Such pairs of RAND and SRES numbers, named response pairs) are downloaded to the HLR, which makes them available to the VLR when needed.

To authenticate, the network sends to a subscriber a RAND number, in a message. MS uses this RAND and the stored Ki as input to the A3 algorithm, to produce SRES number. The value is returned to the VLR and checked for matching. If the two SRES (computed at network level and at MS level) do not match then the user cannot connect to the network. We note that Ki is not sent over the network, for security reason.

- Encryption of any frame is accomplished by using A5 algorithm, using a key named, Kc.

Kc is calculated as output from A8 algorithm, having as input Ki and RAND, mentioned before.

Kc is calculated by AUC and sent to HLR, as well as by the SIM module. The calculated Kc is then exclusive-OR-ed with the 114 bit burst, in both direction.

IMSI Security

There are several defined procedures for managing the Temporary Mobile Subscriber Identity (TMSI). TMSI is a local number and is always associated with Location Area Identification (LAI). VLR maintains a mapping relationship between TSMI and IMSI.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

When a mobile user moves to a region with a new VLR, upon receiving the old TMSI and the LAI, the new VLR queries the old VLR for the IMSI and after reception, it allocates to the MS an new TMSI.

Layer 3 messages

The header of the layer 3 messages consist of:

- protocol discriminator, which contains call control and call related SS messages, mobility management messages and radio resource management messages
- transaction indicator
- message type, identifying the type of message, like: ciphering mode command, channel release, paging request etc.
- information elements.

GSM provides a large variety of messages in order to support the mobile user services.

Communications Procedure

The main steps involved in the communication procedure are:

1. MS finds a proper frequency domain: MS search, finds and locks on a strong frequency, and then finds the proper FCCH.
2. MS locates the proper SCH slot. Using SCH, MS obtains synchronization and training information.
3. MS locate by position BCCH in the frame. By reading BCCH, MS obtains system parameters, such as cell and system values.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4. MS reads RACH to request a stand-alone dedicated channel (SDDCH). The BSC grants this request through AGCH. The response contains also a so called "timing advance information". This information is based on the analysis made by MSC on the signal received from MS, by measuring the delay of this signal. The purpose of the analysis is to not let the MS signal to interfere with signals received from other MSs, on the same physical channel.

5. MS asks for a location update. This is done by comparing the previously stored location, with the information received at event 3. The update is received by a indication in the BCCH signal.

6. Authentication request

7. Authentication response

8. Encryption command

9. Encryption acknowledgement

10. Location update is accomplished and acknowledged.

MS is now ready to originate or accept calls.

Call terminating at MS

The full message flow, when MS receives a call, is:

1. Initial address message

2. Paging operations

3. Channel assignment

4. Authentication operations

5. Ciphering setup

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6. TMSI operations

7. Call setup

8. Call proceeding

9. TCH assignment operations

10. Alerting message

11. Address complete message

12. Connect call

13. Connect acknowledged

14. Answer message

Call originating at MS

The full message flow, when MS originates the call, is:

1. Channel assignment

2. Authentication operations

3. Ciphering setup

4. TMSI operations

5. Call setup

6. Call proceeding

7. TCH assignment operations

8. Initial address message

9. Address complete message

10. Alerting

11. Answer message

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

12. Connect

13. Connect acknowledged

GSM 900, GSM 1800 and its sister standard PCS 1900 are the most widely used second generation cellular technologies in the world.

The North American GSM standard has some differences from its European counterpart, related to the transmission power, and to some additional services like, multiparty service, advice of charge, barring outgoing calls etc.

4.3. Code Division Multiple Access (CDMA) in standard IS-95

IS-95 standard was produced through the Telecommunications Industries Association/Electronic Industries Association (TIA/EIA). It is based on Code Division Multiple Access (CDMA) technology.

CDMA places all users of a cell in the same frequency spectrum at the same time. Each user in a cell is identified by a unique code. At the transmitting side, the code is used to encode the traffic and to spread it across the spectrum. At the receiving side, the code is used to decode the useful information.

CDMA is in commercial use in USA since 1996. The first deployment was in Korea and Hong Kong since 1995.

IS-95 is designed as a dual mode operation, both in AMPS and CDMA environment, therefore working on the AMPS frequency spectrum. For the CDMA operation, multiple AMPS multiple AMPS channels are combined into a 1.288MHz bandwith. This approach

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

allows a quick deployment of the CDMA technology over an already existing technology, the AMPS.

The same 1.288MHz bandwith is used in each cell of the system.

Channel Operations

Two CDMA channels are defined: forward channel and reverse channel. In the AMPS bandwith the two channels are separated by 45Mhz. In the PCS spectrum the two channels are separated by 80MHz.

IS-95 defines two transmission rates:

- rate set 1, operating at 8kbits/sec
- rate set 2 operating at 13kb/s

The two rates have different coding and syntax characteristics. The rate is negotiated between MS and BS, during connection setup.

Channel Structure

1. Forward channels (BS-to-MS) are divided into:

- Control channels:
    - One pilot channel
    - One sync channel
    - 1-7 paging channels
- Traffic Channels

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Each channel is identified by a specific Walsh code, then modulated by a pair of Pseudo-noise sequences, at a fixed chip rate of 1.2288Megachips-per-second (Mcps).

2. Reverse channels (MS-to-BS) are divided into:

- Access channels: for control operations, 1-32 channels, grouped in 20ms frames
- Traffic Channels: for communications

All these channels are identified by a long pseudo-noise sequence. The codes differ in order to discriminate each access channel and each traffic channel.

Since the same bandwith is used in each cell, each BS (base-station) is distinguished by a unique identifier.

Pseudo-noise Code

The code used for encoding and spreading of the traffic is deterministic, but it looks like a random noise for someone not having the correct decoding capabilities.

The pseudo-noise (PN) code has the following properties:

- Balance property: the number of 1s and 0s cannot differ by more than one digit
- The PN sequence must exhibit the run property (Bernoulli sequence):
- Correlation property: if a PN sequence is shifted (left or right) a non-zero number of times, the resulted sequence compared to the original sequence has the number of agreements and disagreements of the two sequences differ by no more than one.

PN sequences generator: PN sequences are generated with a linear feedback shift register, where the feed-back logic is based on a modulo-2 adder. A shift register of length n will produce a PN sequence of $2^n - 1$

IS-95 uses two types of PN codes:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- short code: produced with a shift register of length 15, producing a code of 32,767 bits; as they are generated at 1.2288MHz rate, the codes repeat every 26.67ms. The short codes are used to identify the base station, by generating the same short code among all BS, but with different offsets in time.
- long code: produced with a shift register of 42, producing a code of 2^{42}-1 bits; therefore the codes repeat themselves approximately every 41 days.; for the reverse channel, they are used to identify each channel, and for the forward channel they are used for scrambling.

IS-95 Layers

IS-95 is organized in three logical layers:

- Layer 1 represents the Physical layer
- Layer 2 is subdivided in:
  - Multiplex Sub-layer
  - Upper Sub-layer, which is of three kinds:
    - Primary Traffic
    - Secondary Traffic
    - Signaling: is responsible for the reliable delivery of the Paging and Access Channels
- Layer 3 is responsible for the management of calls, handoffs, power control etc.

Layer 1 Operations

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Functions of forward channels and reverse channels present many similarities as well as some differences, which will be presented.

The following are the major functions at Physical Layer level: Vocoding, Convolutional Coding, Repetition Function, Block Interleaving, Scrambling, Power Control, Orthogonal Spreading, Quadrature Spreading, and Filtering.

Further, we shortly describe each of these functions:

1. Vocoding

A separate standard document, IS-96-A, defines the vocoding standard for CDMA.

On the forward channel, the Vocoder accepts PCM voice traffic and produces a compressed signal. The vocoder is defined to operate at full, half, one quart or one eight rate.

Two types of vocoders are defined:

- rate set 1 vocoder, producing a 8kbs voice stream over a 9.6kbs frame, and
- rate set 2 vocoder, producing 13kbs voice stream over 14.4kbs frame.

2. Convolutional Coding

Convolutional coding gives the possibility to correct some errors at the receiver , therefore giving the possibility to lower the transmission power. Convolutional coding uses trellis coding technique; it is used both for the forward and reverse channels; the used algorithms is being somehow different.

3. Repetition function

It is used to add redundancy to the transmission in order to reduce impairments and to produce a constant input stream to the next process (rate adaptation). With this function the data is repeated a number of times until a rate of 19.2kbs is obtained for rate set 1,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and 28.8kbs for rate set 2.

4. Block interleaving

Interleaving mitigates the effect of Raleigh fading, by preventing that a certain transmission to occupy the same physical channel for a long period of time. By interleaving, the bits are placed in non-contiguous slots, on the channel: every 20 ms, 384 bits are interleaved.

5. Scrambling

The interleaved slots are altered, using a long code PN generator. The input to the generator is the Electronic Serial Number (ESN) of the MS.. The scrambling is performed with a modulo-2 addition operation between the bit-stream and the resulted long code PN. Only the first output of every 64 bits is used for the scrambling. By using the ESN, the user's traffic is uniquely identified from that of other users.. The ESN is made available to the network during the handshaking process, by using the uplink access channel.

6. Power Control and Puncturing

The Base Station (BS) estimates the transmission power from MS, by examining its transmission on the reverse traffic channel, for 1.25ms. After that BS sends its power control message to the MS, over the power control sub-channel, with a "0" indicating an increase of power, and "1" to indicate a decrease in power. The bits are inserted through an operation named "symbol puncturing", by which modulation symbols are replaced by power control bits: for rate set 1, 1 bit and for rate set 2, 2 bits,. are replaced for every 24 bits modulation symbols.

7. Orthogonal Spreading

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The orthogonal spreading takes place by applying a 64 bit orthogonal Walsh code to the bit stream. Each 6symbol block is replaced with the 64-ary Walsh chip. As a result, the 19.2kbs full bit rate (named also chip rate) is spread into 19.2x64=1.2288Mbs.

8. Quadrature Spreading

After the orthogonal spread, each channel is quadrature spread, with a spreading sequence of length 2^{15}, for a 32768 Pseudo-Noise (PN) chip. Two PN sequences are used, named I and Q quadrature.

9. Filtering

Filtering restricts signal frequencies in order to define a pass-band of the signal.

- Reverse Channel Operations:

Vocoding, convolutional coding, symbol repetition, block interleaving are similar to the operations from the forward channels.

Orthogonal Modulation

In this 28.8kbs traffic stream, each 6-symbol block is replaced with a 64-chip Walsh code, producing a 28.8/6*64=307.2kcps rate.

Data Burst Randomization

Data Burst Randomization is implemented in order to reduce interference from other mobile stations. The transmitter divides the 20ms frame into 16 time slots of 1.25ms each, and uses an algorithm to distribute MS's data into all these time slots, throughout the 20ms frame. The result is power spreading and reducing potential interference on the channel.

Direct Sequence Spreading

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Uses a long code mask that is derived from the mobile station's ESN, making the transmission unique to each user.

Quadrature Spreading

Uses a short PN code with a 0 off-set. As a result, the combination of direct sequence spreading and quadrature spreading defines each user in a unique way.

Channel Organization

We can distinguish the following types of channels:

- Pilot Channel: is transmitted continuously by BS, on each CDMA frequency.. It consists of all 0s and is coded with the Walsh code function 0. In order to differentiate between pilot channels from two adjacent cells, the same quadrature spreading code is applied to the traffic, but with different offset values.
- Sync Channel: is used in every cell by the MS to obtain identification information about the cell and information about cell specific PN off-set. The sync message is sent continuously over this channel. The message contains the system time, long code state, information on the paging channel.
- Paging Channels: are optional, but generally used, identified with Walsh codes 1-7. They operate at 4.8kbs or 9.6kbs. Each BS in a cell applies one of the 512 offsets to the signal in order to differentiate between cell sites. The purpose of channel ids to transmit messages from BS to notify MS of an incoming call, and it is usually sent from more than one BS. Also through this channel is sent information about registration, power and about access channels.
- Forward traffic Channels: are identified with a 64 bit Walsh function. When the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

function is applied, a resultant 1.2288Mcps rate is obtained. The repletion rate of the chip is 32,768 chips every 26.66ms.

- Reverse Traffic Channels: User traffic on these channels are identified by a user specific long code sequence, based on ESN.
- Access Channel: is used by MS to communicate with the network. Up to 32 access channels are supported. It operates at 4.8kbps and can be used by more than one mobile station.

Layer 2 Operations

The main function of the Layer 2 is message acknowledgement for a correct exchange of control and signaling messages. They operate at a speed of 4.8kbps and can be used by more than one mobile station.

Layer 3 Operations

The call processing operations at MS include the following states: MS initialization state, MS idle state, System access state, and Mobile Station control on the traffic channel state.

1. Initialization state

Consists of four sub-states: System Determination sub-state: is used to determine which service provider to use (A or B), whether analog or CDMA operations will be employed, if primary or secondary channel will be used

- Pilot Channel Acquisition sub-state: is the operation in which MS acquires the pilot channel of the selected provider (A or B). If the pilot channel is obtained within a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

predetermined period of time, then enters the next substate.

- Sync channel Acquisition sub-state: Ms receives a message on the sync channel containing information about configuration requirements and timing, as:
  - Protocol revision level
  - System Identification (SID)
  - Network Identification (NID), which identifies a network or a part of the cellular system
  - Pilot PN sequence offset (PILOT_PN), used to identify the base station
  - Long code state (LC_STATE), which is used for scrambling on the forward channel and spreading on the reverse channel
  - System time (SYS_TIME)
  - Paging Channel data Rate (PRAT).

After this sub-state, MS enters the next sub-state.

- Timing State sub-state: MS uses the parameter sent through the sync channel, in order to adjust its timing and synchronize its long code timing and system timing, to those of the base station. Also MS sets a value to identify the paging channel, initializes several sequence number parameters that are used for sequencing different types of messages.

After this sub-state, MS enters into Idle State.

2. Idle State

When entering this state, MS sets its code channel to PAGECH and uses the PRAT to set the paging channel data rate.

The MS monitors the paging channel to receive messages, to receive incoming calls etc.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Also in this state MS does the following:

- Paging channel monitoring: there are two modes of operation for this procedure:
    - Non-slotted mode: paging slots are not assigned to MS, which therefore has to monitor all slots
    - Slotted mode: means that slot are specifically assigned to the MS , and will only monitor specific slots on the channel.
- Message Acknowledgement: are used to ensure that the exchange of messages between MS and BS occurs correctly. (it was described as a layer 2 operation)

Registration: When MS is in idle state some timers may have expired. Therefore, MS must initiate idle registration procedures. Through this operation, MS inform the network about its location. There are two categories of idle registration:

- Autonomous Registration: occurs when MS initiates the registration. In this category there are five types of registrations:
    - Power-up registration: when MS first powers-up the system
    - Power-down registration, when MS powers-down
    - Timer-based registration, occurs when initialization timer expired or is disabled
    - Distance-spaced registration: when its distance exceeds a given distance from BS
    - Zone-based Registration when MS registers with a group of cells.
- Non-autonomous registration occurs when it is initiated by the network There are four types of this kind of registration:
    - Parameter-change Registration: MS registers with the system if certain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

parameters of an operation are changed

- Ordered Registration: MS must register this way when it is ordered so by the system
- Implicit Registration: is an ongoing registration performed by BS, by receiving ongoing information from MS (page response message, origination message).
- Traffic Channel Registration occurs when a traffic channel is assigned to MS.

- Idle Handoff:

There are four type of handoffs:

- hard handoff
- soft handoff
- softer handoff
- idle handoff

The first three are performed when the mobile station is on a traffic channel. The idle handoff occurs when MS moves from one BS coverage to another BS coverage and it is in an idle state. The idle handoff procedure requires the MS to operate in a non-slotted mode, until it has received at least one valid message on the new paging channel.

- Response to overhead information: MS may receive several types of overhead messages on the paging channel:
  - System parameter messages
  - Access parameter messages
  - Neighbor list message

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- CDMA channel list message
- Extended system parameter message
- Global service redirection message

The purpose of these messages is to provide MS with system parameters. MS responds to these messages depending on the type of message.

- Page Match: Three types of messages are transmitted over the paging channel:
  - Page message
  - Slotted page message
  - General page message.

The MS checks the matching of certain page parameters with that of the network. The most important of these is the match MIN1 and MIN2 and the MS storage of IMSI_S1 and IMSI_S2.

- Order and message processing: If MS performed successfully a address matching, then has to process the message; otherwise the message is ignored. The following messages can be received and are processed, during this operation:
- Abbreviated alert order
- Audit order
- Authentication challenge message
- Base station acknowledgement order
- Base station challenge confirmation order
- Channel assignment message
- Data burst message
- Feature notification message

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Local order control order message
- Lock until power cycled order message
- Maintenance required order
- Registration accepted order
- Registration rejected order
- Registration request order
- Service redirection
- SSD update
- Unlock order
- Origination operation: is executed when MS is activated by the user, to begin a call. When this occurs, MS enters the update overhead information substate or the system access state.
- Data burst transmission: is an optional procedure. If supported, MS reacts to the user by directing it to transmit a data burst message. Then MS enters the "Update overhead information sub-state" of the "System access state".
- Power-down operation: is executed when the user of MS directs the MS to power down.

3. System Access state

In this state MS sends messages to BS on the access channels and receives messages on the paging channel. During this state, the following sub-states are possible:

- Update overhead information substate: MS monitors the paging channel, until it has a current set of overhead messages

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- MS origination attempt sub-state: MS sends an Origination Message to BS
- Page response sub-state: MS sends a Page Response Message to BS
- MS order/message response sub-state: MS sends a response to a message received from BS.
- Registration access sub-state: MS sends a Registration Message to BS.
- MS message transmission sub-state: MS sends a data burst message to BS.

4. Handoff

CDMA technology has two kind of handoff:

1. Soft Handoff: is the operation in which the MS is simultaneously communicating with several BS and makes decision about which BS to use, based on the signal strength. Depending on the movement of the MS, a certain Bs could be removed from the active set of BSs, when the signal strength goes below a given threshold. The entries in the active set may fluctuate in and out of the set, depending mainly on terrain, reflections and other temporary impairment of the signal from the BS.

2. Hard Handoff: is the operation in which MS is communicating with a single BS at a time. MS relinquishes resources in the old BS and then tunes to a new BS. Therefore hard handoff is also named "brake before make" as opposed to soft handoff, which is named "make before brake".

Hard handoff is usually made in the following three situations:

- moving between BSs served by different switches
- BSs belong to different vendors
- BSs are supporting different air interfaces (like CDMA to AMPS).

Disadvantages of hard handoff are:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- requires more power
- may be discernable to the user.

Communications Procedure

Originating a call

Initial start-up operations for call origination from MS:

1, MS seeks for and acquires the pilot channel

2.MS receives the sync channel

3.MS adjusts its timing parameters through the sync channel

4. MS receives administrative and operations information through the paging channel

5.MS creates an origination message to place a call and sends the origination message through the access channel to BS

6.BS analyzes this message and creates the channel assignment message

7.This message is sent to MS through the paging channel

8.If the message is successfully received, MS sends on the reverse channel the preamble pattern

9.After receiving the preamble pattern, BS creates an acknowledgement order message and sends it over the forward traffic channel

2. This prompts MS to begin sending null channel data on the reverse channel
3. BS creates a "service option order" message and sends it over the forward traffic channel.

This completes the initial start-up operation for a call origination from MS.

Receiving a call

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. MS seeks for, and acquires the pilot channel

2. MS receives the sync channel

3. MS adjusts its timing parameters through the sync channel

4. MS receives administrative and operations information through the paging channel

5. BS sends an alert with information message on the forward traffic channel

6. MS starts ringing

7. If user answers the call, MS creates the "connect order" message and sends this message to Bs over the reverse traffic channel

8. The same time Ms removes ringing.

9. Upon correct reception of the "connect order" message, Bs starts sending the user traffic to MS.

4.4. Internetworking standard IS-41

IS-41 is a North American standard that works on the network side of the wireless networks.

IS-41 specifies over 100 operations and is related to the transport of messages between MSC and its corresponding databases. It allows the inter-working of different vendors' equipment.

IS-41 was initially published in 1988. Since then, it undergone several revisions (presently Revision D is in use).

Functional planes

The most widely used version, IS-41-C, has the following functional entities:

- Functional Overview

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Intersystem Handoff Information Flow
- Automatic roaming Information Flow
- Operations, Administration and Maintenance Information Flows and Procedures
- Signaling Protocols
- Signaling Procedures

This standard covers only the interfaces with: Home Location Register (HLR), Visitor Location Register (VLR) and Authentication Center (AC).

Further we describe standard IS-41-C, organized around the OSI model. The first three layers could be those of the transport technology used (SS7, or X.25, Frame Relay, ATM etc.).At Application Layer level the following sub-layers are used:

- Mobile Application Part (MAP)
- Application Service Element (ASE), known also as Association Control Service Element (ASCE): used to bind two applications together, allowing the negotiation of different services and operations, dependent upon a specific network or switch.
- Remote Operation Service Element (ROSE) was included in Transaction Capabilities Application Part (TCAP). TCAP also contains X.208, which is a compiler independent language, using Abstract Syntax Notation (ASN.1) to code structures. TCAP also contains X.209 which defines the syntax of IS-41-C messages. TCAP is based on sending a request for an operation to a server and conveying the results of that operation to the client.

Security Operations

The authentication operations imply the exchange of between the network and MS to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ensure that the user of the MS has access right to the network. The idea is to use a secret key named Shared Secret Data (SSD), that is known only to the network and the Mobile Station (MS). A successful authentication means that:

- MS and the network have identical copies of the SSD
- Use the SSD to generate other values exchanged in order to authenticate the MS and the network
- Use the SSD to encrypt the voice or data traffic.

IS-41-C provides five authentication/privacy operations:

- Authentication of Mobile Station (MS) registration
- Unique challenge-response procedures
- Authentication of Mobile Station (MS) originating a call
- Authentication of a call to a terminating Mobile Station (MS)
- Updating of the Secret Shared Data (SSD).

Handoff between MSCs

A serving MSC can query its adjacent MSCs to determine if MS should be relocated to another serving system. In order to do that, the following sequence of events will happen:

1. The serving MSC sends to its neighboring MSCs the request of a measurement level on a specific channel.
2. The adjacent MSCs respond with the handoff-measure-request response, containing the identifier of the responding cell and the signal strength, received on the specific channel.
3. The serving MSC examines each response to determine if a handoff is/or is not appropriate.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Due to the widespread use of the GSM standard, internetwork services gateways (ISG) were defined, for interworking with the IS-41 standard. When a GSM user roams on a IS-41 network, ISG operates as a HLR to an IS-41 MSC/VLR, and also it acts as an IS-41 VLR to a GSM HLR.

ISG participates in call establishment handling functions as registration, authentication, location updates, etc.

4.5. BS to MSC communication standard IS-634

IS-634 is a North American specification that defines the operation between Base Station (BS) and Mobile Switching Center (MSC), for the 800MHz air interface.

IS-634 describes the following partitions, named functional planes:

- Transmission Facilities Management: manages the transmission media
- Radio resources Management: supports the link between BS and MSC as well as handoffs
- Mobility Management: Manages databases and subscriber location data
- Call Processing: Manages call control and communications services.

IS-634 supports the following air interfaces, by defining physical layers, channel structures, message types and operating procedures for:

- EIA-553 for AMPS
- IS-54-B for dual mode MS
- IS-91 for BS compatibility
- IS-95 for BS compatibility for wideband spread spectrum.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bibliography

4.1. Black, Uyless. "Second Generation Mobile and Wireless Networks". Prentice Hall, Upper Saddle River, NJ. 1999.

4.2. "Cellular Intersystem Operations (Rev. C)". Technical Report IS-41. EIA/TIA. 1995.

4.3. Lin, Yi-Bing, Imrich Chlamtac. "Wireless and mobile Networks Architecture". John Willey and Sons, New York, NY. 2001

4.4 Papadimitriou, G. "Wireless Networks". Wiley and Sons. December 2002

Prasad, R. "CDMA for Wireless Personal Communications". Artech House. Norwood, MA. 1996.

4.5. Redl, S., M. Weber, M. Oliphant. "Introduction to GSM". Artech House. Norwood MA. 1995.

4.6 Snyder, Randall A. "Wireless Telecommunications Networking with ANSI-41". McGraw Hill. 2001.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 5. Third Generation Wireless Networks

3rd generation wireless networks represent an important step ahead for providing communications to "everyone, everywhere at any time". To achieve this goal, it has to provide the necessary functionalities to different type of users, as:

- cellular users, with high voice quality and good coverage
- users of private network systems, eventually with group or broadcast calls
- paging users with small terminal and good coverage
- cordless users, in homes or offices, with both voice and high speed data communication
- satellite users with worldwide coverage
- users in airplanes (having at this time very limited availability)
- data users with high mobility or remotely located
- fixed/mobile convergence, with similar set of services and using the same phone number.

5.1. Key Features

The principal goals of this system are:

- Common provisioning for all subscribers, either fixed or mobile
- Universal subscriber identifier to support portable telephone numbers
- ATM-based switching architecture
- Services to al subscribers, either fixed or mobile, by common call-processing procedures

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The following key features and objectives are envisioned:

1. Global System:

- A global standard with commonality of design, incorporating a variety of systems and terminal types
- Worldwide market-place and compatible equipment
- Worldwide common frequency-band and roaming

2. New Services and Capabilities:

- New capabilities enabling new, advanced audio, video and data communications services
- High service quality and integrity (comparable to the fixed network)
- Improved spectrum efficiency at lower cost
- Provision of "bandwith-on-demand", supporting a large range of data rates, from low paging messages to voice to high video transfer
- Support of asymmetrical data capabilities (high rate for download, low rate for upload)
- Improved security and ease of operation
- Distributed control of service creation and service profile management
- Coherent systems management

3. Evolution and migration:

- Flexibility for evolution and migration of systems, both pre-IMTS as well as within
- Compatibility of services within IMTS and with PSTN
- Provision for easy and continuous evolution of services and access to services of fixed network

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Open architecture which will allow easy introduction of new technologies and features

4. Flexibility for multi-environment capabilities:

- Maximum level of interworking between networks of different types
- Integrated satellite-terrestrial networks
- Provision of services by more than one network, in any coverage area
- Provision of services over a variety of population densities and coverage areas
- Provision of services for both fixed and mobile users
- A wider range of operating environments (including aeronautical or naval)
- Modular and scalable structure to support evolution and emerging needs
- Use of adaptive software-downloadable terminals to support multiband and multi-environment capabilities
- Capability to select and negotiate bandwith, transmission quality, delay etc. according to instantaneous capability of the radio channel
- Better use of radio spectrum, in order to provide services, at acceptable cost.

Four different kind of 3G specifications did emerge at this time. All of these use wideband CDMA(WCDMA) and have similar high speed capabilities. Wideband CDMA does only mean just a wider bandwith that 2G systems have.

5.2. Phase 2+G

For a smooth migration to 3G systems, an intermediary 2+ generation system step was conceived.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The key features to be added to existing 2nd generation wireless networks, to transition toward 3rd generation, are:

- Advanced speech call items: are defined as group and broadcast calls.

  Broadcast calls allow many users to listen to a single radio channel in a cell, therefore the call can be available in many cells.

  Group calls allow users to respond after the originator finished speaking.

- General Packet Radio Service (GPRS) allows packet switched communications. It is generally seen as an important step toward UMTS-2000, that will become predominantly packet based.
- Higher circuit switched data rate, named High Speed Circuit Switched Data (HSCSD) will provide higher rate data transmission by concatenating time slots.
- Intelligent network capabilities, named Customized Applications for Mobile-network Enhanced Logic (CAMEL), will allow subscribers to use operator specific services even during roaming to an other network.
- Enhanced full-rate codec (EFR) will offer higher speech quality, without reducing the coder rate.
- Combined cellular/cordless handset, will allow the use of handset as a cordless phone in-house and as a cellular phone if outdoor.

Mobile system providers and carriers have developed 2G+ technology that is packet-based and increases the data communication speeds to as high as 384kbps. These 2G+ systems are based on the following technologies:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

wideband mobile data networks. This circuit-switched technology improves the data rates up to 57.6kbps by introducing 14.4 kbps data coding and by aggregating 4 radio channels timeslots of 14.4 kbps.

General Packet Radio Service (GPRS): is an intermediate step that is designed to allow the GSM world to implement a full range of Internet services without waiting for the deployment of full-scale 3G wireless systems. GPRS technology is packet-based and designed to work in parallel with the 2nd generation systems like: GSM, PDC and TDMA. GPRS uses a multiple of the 1 to 8 radio channel timeslots in the 200kHz-frequency band allocated for a carrier frequency to enable data speeds of up to 115kbps. The data is packetized and transported over Public Land Mobile Networks (PLMN) using an IP backbone so that mobile users can access services on the Internet, such as SMTP/POP-based e-mail, ftp and HTTP-based Web services.

3. Enhanced Data Rates for Global Evolution (EDGE) technologies: is a standard that has been specified to enhance the throughput per timeslot for both HSCSD and GPRS. The enhancement of HSCSD is called ECSD, whereas the enhancement of GPRS is called EGPRS. In ECSD, the maximum data rate will not increase from 64 kbps due to the restrictions in the A interface, but the data rate per timeslot will triple. In EGPRS, the data rate per timeslot will triple and the peak throughput, including all eight timeslots in the radio interface, will be at least 384 kbps.

GPRS networks consist of an IP-based Public Mobile Land Network (PLMN), Base Station

Services (BSS), Mobile handsets (MS), and Mobile Switching Centers (MSC) for circuit-switched network access and databases.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Serving GPRS Support Nodes (SGSN) and Gateway GPRS Support Nodes (GGSN) make up the PLMN. Roaming is accommodated through multiple PLMNs.

SGSN and GGSN interfaces will be used to retrieve the mobile user's profiles from the Home Location Register (HLR), to facilitate call completion. GGSN provides the connection to external Packet Data Network (PDN), e.g. an Internet backbone or an X.25 network.

The BSS consists of Base Transceiver Stations and Base Station Controllers. The Base Transceiver Station (BTS) receives and transmits over the air interfaces (CDMA, TDMA), providing wireless voice and data connectivity to the mobile handsets. Base Station Controllers (BSC) route the data calls to the packet-switched PLMN over a Frame Relay (FR) link and the voice calls to the Mobile Switching Center (MSC).

MSC switches the voice calls to circuit-switched PLMN network such as PSTN and ISDN. MSC accommodates the Visitor Location Register (VLR) to store the roaming subscriber information. The reverse process happens at the destination PLMN and the destination BSS.

On the data side, the BSC routes the data calls to the SGSN, and then the data is switched to the external PDN through the GGSN or to another mobile subscriber.

GRPS Protocols

The following is a brief description of each protocol layer in the GPRS network infrastructure:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. Sub-Network Dependent Convergence Protocol (SNDCP): protocol that maps a network level protocol, such as IP or X.25, to the underlying logical link control. SNDCP also provides other functions such as compression, segmentation and multiplexing of network-layer messages to a single virtual connection.

2. Logical Link Control (LLC): a data link layer protocol for GPRS which functions similar to Link Access Protocol – D (LAPD). This layer assures the reliable transfer of user data across a wireless network.

3. Base Station System GPRS Protocol (BSSGP): processes routing and quality of service (QoS) information for the BSS. BSSGP uses the Frame Relay Q.922 core protocol as its transport mechanism.

4. GPRS Tunnel Protocol (GTP): protocol that tunnels the protocol data units through the IP backbone by adding routing information. GTP operates on top of TCP/UDP over IP.

5. GPRS Mobility Management (GMM/SM): is a protocol that operates in the signaling plane of GPRS, handles mobility issues such as roaming, authentication, selection of encryption algorithms and maintains PDP context.

6. Network Service: is a protocol that manages the convergence sub-layer that operates between BSSGP and the Frame Relay Q.922 Core by mapping BSSGP's service requests to the appropriate Frame Relay services.

7. BSSAP+: protocol that enables paging for voice connections from MSC via SGSN, thus optimizing paging for mobile subscribers. BSSAP+ is also responsible for location and routing updates as well as mobile station alerting.

8. SCCP, MTP3, MTP2 are protocols used to support Mobile Application Part (MAP) and BSSAP+ in circuit switched PLMNs.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

9. Mobile Application Part (MAP): supports signaling between SGSN/GGSN and HLR/AuC/EIR.

5.3. 3rd Generation Wireless Systems

The 3rd generation wireless cellular technology is a federation of systems conceived to provide wireless access to the global telecommunications infrastructure.

ITU named third generation networks IMT-2000 and ETSI (in Europe) as UMTS.

The main goals of the third generation system are wide and solutions will be different. Different needs and different backward compatibility may result in several different variants of IMT-2000 system.

As a federation of standards, IMT-2000 include research and standardization efforts from USA, Europe, Japan, China and Korea:. 13 CDMA based FDD proposals and 3 CDMA-based TDD proposals were made. All these proposals were grouped in 4 type of systems.

The 4 types of 3G systems are known:

1) UMTS FDD and TDD modes developed by 3GPP (3rd generation Partnership Project, a consortium of standardization bodies, including: ARIB, CWTS, ETSI, T1, TTA, and TTC)
2) Cdma2000 developed by 3GPP2 (3rd generation Partnership Project 2, defines itself as a collaborative third generation (3G) telecommunications specifications-setting project, representing North American and Asian interests)
3) WCDMA in Japan, containing ARIB specifications
4) TD-SCDMA mainly in China, supported by TD-SCDMA Forum

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Of these systems, W-CDMA European standard and cdma-2000 American standard emerged as the main contenders for IMT-2000. Even if both are CDMA-based, they have many different features, as:

- W-CDMA uses dedicated TDM pilot signal which reduces overall pilot power and does not need base station timing synchronization, and Cdma-2000 uses CDM pilot and base station synchronization resulting in decreased latency and reduced chance of dropping calls.
- W-CDMA is a Direct Sequencing frequency division duplex mode (DS-FDD), and cdma-2000 is a multicarrier FDD mode (MC-FDD).

Beside that, a Time Division Duplex Mode (TDD) is considered, mainly for indoor cordless communications.

For roaming, all radio access modes should fully support both ANSI-41 and GSM-MAP.

For the satellite component of IMT-2000 proposals have been submitted to ITU-R from ESA, ICO and INMARSAT.

Frequency allocation

World Administrative Radio Conference (WARC) allocated spectrum to 3rd generation since 1992. Europe and Japan followed these recommendations but in USA, FCC already allocated a good part of the spectrum for 2nd generation PCS systems.

The main objectives of the IMT-2000 air interface are:

- Full coverage and mobility at 384Kbs
- Limited coverage and mobility at 2Mbs
- High spectrum efficiency (compared to 2nd generation systems)
- High flexibility in introducing new services

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Backward compatibility with existing 2nd generation systems

3G wireless network general architecture

Between other goals, 3G wireless technology will solve the convergence of various 2G wireless telecommunications systems into a single global system that includes both terrestrial and satellite components. One of the most important aspects of 3G wireless technology is its ability to unify existing cellular standards, such as CDMA, GSM, and TDMA, under one umbrella.

The following three air interface modes accomplish this result: Wideband-CDMA, CDMA2000 and the Universal Wireless Communication (UWC-136) interfaces.

1. Wideband CDMA (W-CDMA) is compatible with the current 2G GSM networks used mainly in Europe and some Asian countries. W-CDMA will require bandwidth of between 5Mhz and 10 MHz, making it a suitable platform for higher capacity applications. It can be implemented over existing GSM, TDMA(IS-36) and IS95 networks. Subscribers are likely to access 3G wireless services initially via dual band terminal devices. W-CDMA networks will be used for high-capacity applications and 2G digital wireless systems will be used for voice calls.

2. CDMA2000 is backward compatible with the second generation CDMA IS-95 standard, predominantly used in US.

3. Universal Wireless Communications – UWC-136, also called IS-136HS, was proposed by the TIA and designed to comply with ANSI-136, the North American TDMA standard.

3G wireless networks consist of a Core Network and a Radio Access Network (RAN).

1. The core network consists of :

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- a packet-switched domain, which includes SGSNs and GGSNs, (providing the same functionality that they provide in a GPRS system), and

- a circuit-switched domain, which includes 3G MSC for switching of voice calls.

Charging for services and access is done through the Charging Gateway Function (CGF), which is also part of the core network.

2. RAN functionality is independent from the core network functionality. The Radio Access Network provides a core network technology independent access for mobile terminals to different types of core networks and network services. Either core network domain can access any appropriate RAN service; e.g. it should be possible to access a "speech" radio access bearer from the packet-switched domain.

The Radio Access Network consists of new network elements, known as Node B and Radio Network Controllers (RNCs).

- Node B is comparable to the Base Transceiver Station in 2G wireless networks.

- RNC replaces the Base Station Controller. It provides the radio resource management, handover control and support for the connections to circuit-switched and packet-switched domains.

- The interconnection of the network elements in RAN and between RAN and core network is over Iub, Iur and Iu interfaces using ATM as a layer 2 switching technology. Data services run from the terminal device over IP and uses ATM as a reliable transport with QoS. Voice is embedded into ATM from the edge of the network (Node B) and is transported over ATM out of the RNC.

The Iu interface is split into 2 parts: circuit-switched and packet-switched.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Iu interface is based on ATM with voice traffic embedded on virtual circuits, using AAL2 technology and IP-over-ATM for data traffic, using AAL5 technology. These traffic types are switched independently to either SGSN for data, or MSC for voice.

The general architecture of a 3rd generation wireless network is presented in the figure below:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

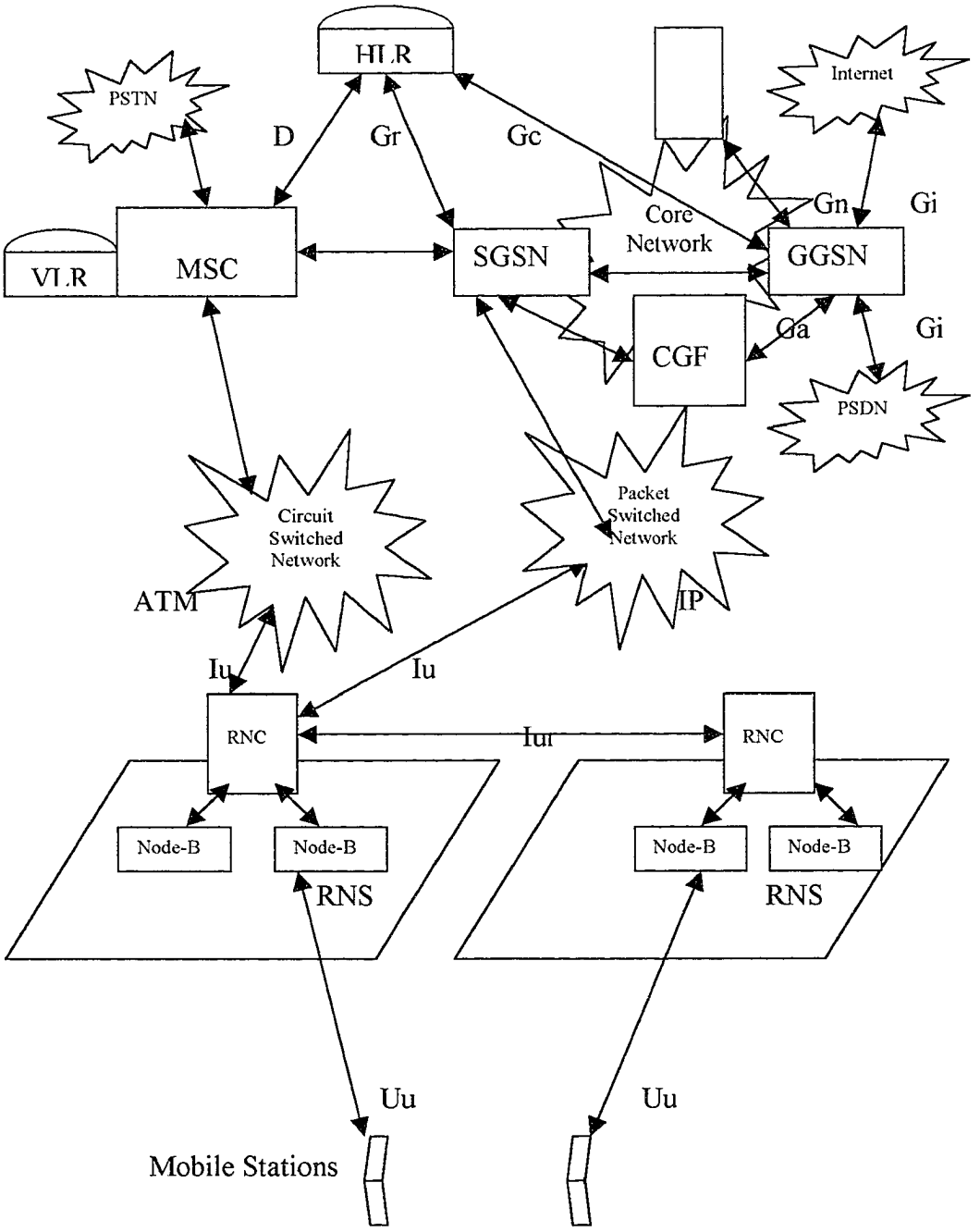


Figure 5.1. 3rd Generation Wireless Network - General Architecture

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The following is a brief description of each protocol layer in a 3G wireless network infrastructure:

1. Global Mobility Management (GMM): protocol that includes attach, detach, security, and routing area update functionality.

2. Node B Application Part (NBAP): provides procedures for paging distribution, broadcast system information and management of dedicated and logical resources.

3. Packet Data Convergence Protocol (PDCP): maps higher level characteristics onto the characteristics of the underlying radio-interface protocols. PDCP also provides protocol transparency for higher layer protocols.

4. Radio Link Control (RLC): provides a logical link control over the radio interface.

5. Medium Access Control (MAC): controls the access signaling (request and grant) procedures for the radio channel.

6. Radio resource Control (RRC): manages the allocation and maintenance of radio communication paths.

7. Radio Access Network Application Protocol (RANAP): encapsulates higher layer signaling, manages the signaling and GTP connections between RNC and 3G-SGSN, and signaling and circuit-switched connections between RNC and 3G MSC.

8. Radio Network Service Application Part (RNSAP): provides the communication between RNCs.

9. GPRS Tunnel Protocol (GTP): tunnels the protocol data units through the IP backbone by adding routing information. GTP operates on top of TCP/UDP over IP.

10. Mobile Application Part (MAP): supports signaling between SGSN/GGSN and HLR/AuC/EIR.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

11. AAL2 Signaling (Q.2630.1, Q.2150.1, Q.2150.2, AAL2 SSSAR, and AAL2 CPS): protocols suite used to transfer voice over ATM backbone using AAL2.

12. Sigtran (SCTP, M3UA): is a protocol suite, used to transfer SCN signaling protocols over IP network.

5.4. Universal Mobile Telecommunications System (UMTS)

The Universal Mobile Telecommunications System (UMTS) is one of the 3G wireless network systems developed within the ITU's IMT-2000 framework, promoted by the European standard organization, ETSI.

The worldwide coverage will be provided by a combination of cell sizes ranging from pico-cells, used in in-door communications, to global cells provided by satellite communications.

UMTS uses the following frequencies and channel spacing:

-1900-1920 and 2010-2025 MHz Time Division Duplex (TDD, TD/CDMA) Unpaired, channel spacing is 5 MHz and raster is 200 kHz. Tx and Rx are not separated in frequency

-1920-1980 and 2110-2170 MHz Frequency Division Duplex (FDD, W-CDMA) Paired uplink and downlink, channel spacing is 5 MHz and raster is 200 kHz. An Operator needs 3 - 4 channels (2x15 or 2x20 MHz) to be able to build a high-speed, high-capacity network.

-1980-2010 and 2170-2200 MHz Satellite uplink and downlink

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Channels

UMTS air interface logical channels are:

- Broadcast Control Channel (BCCH)

- Paging Control Channel (PCCH)

- Dedicated Control Channel (DCCH)

- Common Control Channel (CCCH)

- Dedicated Traffic Channel (DTCH)

- Common Traffic Channel (CTCH)

UMTS will offer the following data rates of the services:

-2.048Mb/s for micro-cell and pico-cell applications.

- 384kb/s for medium size cells.

- 144kb/s and 64kb/s for large cell applications.

- 14.4kb/s for continuous low speed data applications in very large cells.

- 12.2kb/s for speech (4.75kb/s - 12.2kb/s)

Modulation Scheme

UMTS WCDMA modulation is Quadrature Phase Shift Keying (QPSK) with Root-raised cosine pulse shaping filters (roll off = 0.22)

UMTS Mobile Multimedia services

Several important group of services are envisioned, and implemented through mobile multimedia applications:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. For General Fun, containing services like: world wide web access, video-post card, snapshots, text, picture and multimedia messaging, datacast, personalization applications (ring tone, screen saver, desk top), jukebox, virtual companion / pet .

2. For Work, containing services like: Rich call with image and data stream, IP telephony, Business-to business communications, information exchange, personal information manager, dairy, scheduler, note pad, 2-way video conferencing, directory services, travel assistance, work group, FTP, instant voicemail, color fax.

3. For Media Communications, containing services like: Push newspaper and magazines, advertising, classified sections, soft-radio, soft-TV.

4. For Shopping, containing services like: E-commerce, e-cash, e-wallet, credit card access, telebanking, automatic transaction, e-auction

4. For Entertainment, containing services like: news, stock market, sports, games, lottery, gambling, music, video, concerts, adult content

5. For Education, containing services like: Online libraries, search engines, remote attendance, field research.

6. For Tele-Security, containing services like: Remote surveillance, location tracking, emergency use

7. For Health, containing services like: Telemedicine, remote diagnose and heath monitoring

8. For Automation: Home automation, traffic telematics, telemetry.

9. For Travel, containing services like: location sensitive information and guidance, e-tour, location awareness, time tables, e-ticketing .

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMTS Architecture

A UMTS network consist of three interacting domains:

- Core Network (CN),
- UMTS Terrestrial Radio Access Network (UTRAN) and
- User Equipment (UE).

1. Core Network

The main function of the core network is to provide switching, routing and transit for user traffic. Core network also contains the databases and network management functions.

The basic Core Network architecture for UMTS is based on GSM network with GPRS. All equipment has to be modified for UMTS operation and services.

The Core Network is divided in circuit switched and packet switched domains. Circuit switched elements are Mobile services Switching Centre (MSC), Visitor location register (VLR), and Gateway MSC and packet switched elements are Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). Some network elements, like EIR HLR and AUC, are shared by both domains.

The Asynchronous Transfer Mode (ATM) is defined for UMTS core transmission. ATM Adaptation Layer type 2 (AAL2) handles circuit switched connection and packet connection protocol AAL5 is designed for data delivery.

The architecture of the Core Network may change when new services and features are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

introduced. Number Portability Data-Base (NPDB) will be used to enable user to change the network while keeping their old phone number. Gateway Location Register (GLR) may be used to optimize the subscriber handling between network boundaries. MSC, VLR and SGSN can merge to become a UMTS MSC.

2. Radio Access (UTRAN)

Universal Mobile Telecommunications System Terrestrial Radio Access Network (UTRAN) provides the air interface access method for User Equipment. In the UTRAN terminology, the Base Station is named Node-B and the control equipment for Node-B is named Radio Network Controller (RNC).

Wide band CDMA technology was selected for UTRAN air interface. UMTS WCDMA is a Direct Sequence CDMA system where user data is multiplied with quasi-random bits derived from WCDMA Spreading codes. In UMTS, in addition to channelization, codes are used for synchronization and scrambling. WCDMA has two basic modes of operation: Frequency Division Duplex (FDD) and Time Division Duplex (TDD).

- The functions of Node-B are:

1. Air interface Transmission / Reception
2. Modulation / Demodulation
4. CDMA Physical Channel coding
4. Micro Diversity
5. Error Handling

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6. Closed loop power control

- The functions of RNC are:

1. Radio Resource Control
2. Admission Control
3. Channel Allocation
4. Power Control Settings
5. Handover Control
6. Macro Diversity
7. Ciphering
8. Segmentation / Reassembly
9. Broadcast Signaling
10. Open Loop Power Control

3. User Equipment (UE)

It is necessary for a network to know the approximate location in order to be able to page an user equipment. In the UMTS standard several type of system areas are defined. Here is the list of system areas from largest to smallest.

1. UMTS systems (including satellite)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. Public Land Mobile Network (PLMN)

3. MSC/VLR or SGSN

4. Location Area

5. Routing Area (PS domain)

6. UTRAN Registration Area (PS domain)

7. Cell

8. Sub cell

The UMTS standard does not restrict in any way the functionality of the User Equipment. Terminals work as an air interface counterpart for Node-B and have many different types of identities. Most of these UMTS identity types are taken directly from GSM specifications.

1. International Mobile Subscriber Identity (IMSI)

2. Temporary Mobile Subscriber Identity (TMSI)

3. Packet Temporary Mobile Subscriber Identity (P-TMSI)

4. Temporary Logical Link Identity (TLLI)

5. Mobile station ISDN (MSISDN)

6. International Mobile Station Equipment Identity (IMEI)

7. International Mobile Station Equipment Identity and Software Number (IMEISV)

An UMTS Mobile Station (MS) can operate in one of the following three modes:

1. PS/CS mode of operation: The MS is attached to both the PS domain and CS domain, and the MS is capable of simultaneously operating PS services and CS services.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. PS mode of operation: The MS is attached to the PS domain only and may only operate services of the PS domain. However, this does not prevent CS-like services to be offered over the PS domain (like VoIP).

3. CS mode of operation: The MS is attached to the CS domain only and may only operate services of the CS domain.

An UMTS IC card has the same physical characteristics as GSM SIM card.

The main functions of the IC card are:

1. Support of one User Service Identity Module (USIM) application (optionally more that one)

2 Support of one or more user profile on the USIM

3. Update USIM specific information over the air

4. Security functions

5. User authentication

6. Optional inclusion of payment methods

7. Optional secure downloading of new applications

5.5. Cdma2000

Third Generation Partnership Project 2 (3GPP2) is a 3G wireless standardization organization, and its CDMA2000 is mainly used in North America and some parts of Asia. 3GPP2 was born outside of the ITU's IMT2000, covering high speed, broadband, and Internet Protocol (IP)-based mobile systems featuring network-to-network interconnection, feature/service transparency, global roaming and seamless services

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

independent of location. IMT-2000 is intended to bring high-quality mobile multimedia communications, by increasing the speed and ease of wireless communications, and by carrying data communications through the wireless.

3GPP2 is a collaborative effort between five officially recognized SDOs:

ARIB - Association of Radio Industries and Businesses (Japan)

CWTS - China Wireless Telecommunication Standards Group (China)

TIA - Telecommunications Industry Association (North America)

TTA - Telecommunications Technology Association (Korea)

TTC - Telecommunications Technology Committee (Japan)

These SDOs are known as the Project's Organizational Partners (OPs). 3GPP2 requires that a participating individual member company be affiliated with at least one of the Organizational Partners.

In addition, 3GPP2 is cooperating with three Market Representation Partners (MRPs), who offer market advice to 3GPP2 and bring a consensus view of market requirements (e.g., services, features and functionality), which are of interest to the project: The CDMA Development Group (CDG), The Mobile Wireless Internet Forum (MWIF), Wireless Multimedia Forum (WMF).

The cdma2000 Family of Standards

The cdma2000 family of standards includes a core air interface, a minimum performance, and three service standards.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The cdma2000 air interface standards specify a spread spectrum radio interface that uses Code Division Multiple Access (CDMA) technology to meet the requirements for 3G wireless communication systems.

In addition, the family includes a standard that specifies analog operation, to support dual-mode mobile stations and base stations.

Physical Channels

Physical channels are represented by upper case abbreviations. As in the case of logical channels, the first letters in the names of the channels indicate the direction of the channel (i.e., forward or reverse) and is followed by a hyphen.

The table below shows the names and meanings of all the physical channels designated in cdma2000.

Channel Name	Physical Channel
F/R-FCH	Forward/Reverse Fundamental Channel
F/R-DCCH	Forward/Reverse Dedicated Control Channel
F/R-SCCH	Forward/Reverse Supplemental Code Channel
F/R-SCH	Forward/Reverse Supplemental Channel
F-PCH	Paging Channel
F-QPCH	Quick Paging Channel
R-ACH	Access Channel
F/R-CCCH	Forward/Reverse Common Control Channel
F/R-PICH	Forward/Reverse Pilot Channel
F-APICH	Dedicated Auxiliary Pilot Channel

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

F-TDPICH	Transmit Diversity Pilot Channel
F-ATDPICH	Auxiliary Transmit Diversity Pilot Channel
F-SYNCH	Sync Channel
F-CPCCH	Common Power Control Channel
F-CACH	Common Assignment Channel
R-EACH	Enhanced Access Channel
F-BCCH	Broadcast Control Channel

(The notations "F/R" and "Forward/Reverse" represent two different physical channels (i.e., one forward channel and one reverse channel)

Cdma2000 general architecture

The cdma2000 family of standards reached a common architecture by specifying different layers from different standards.

The following layers are standardized:

- the physical layer
- the medium access control (MAC) layer
- the link access control (LAC) layer
- upper layer signaling.

The figure 5.2 shows the general architecture of the cdma2000 system:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

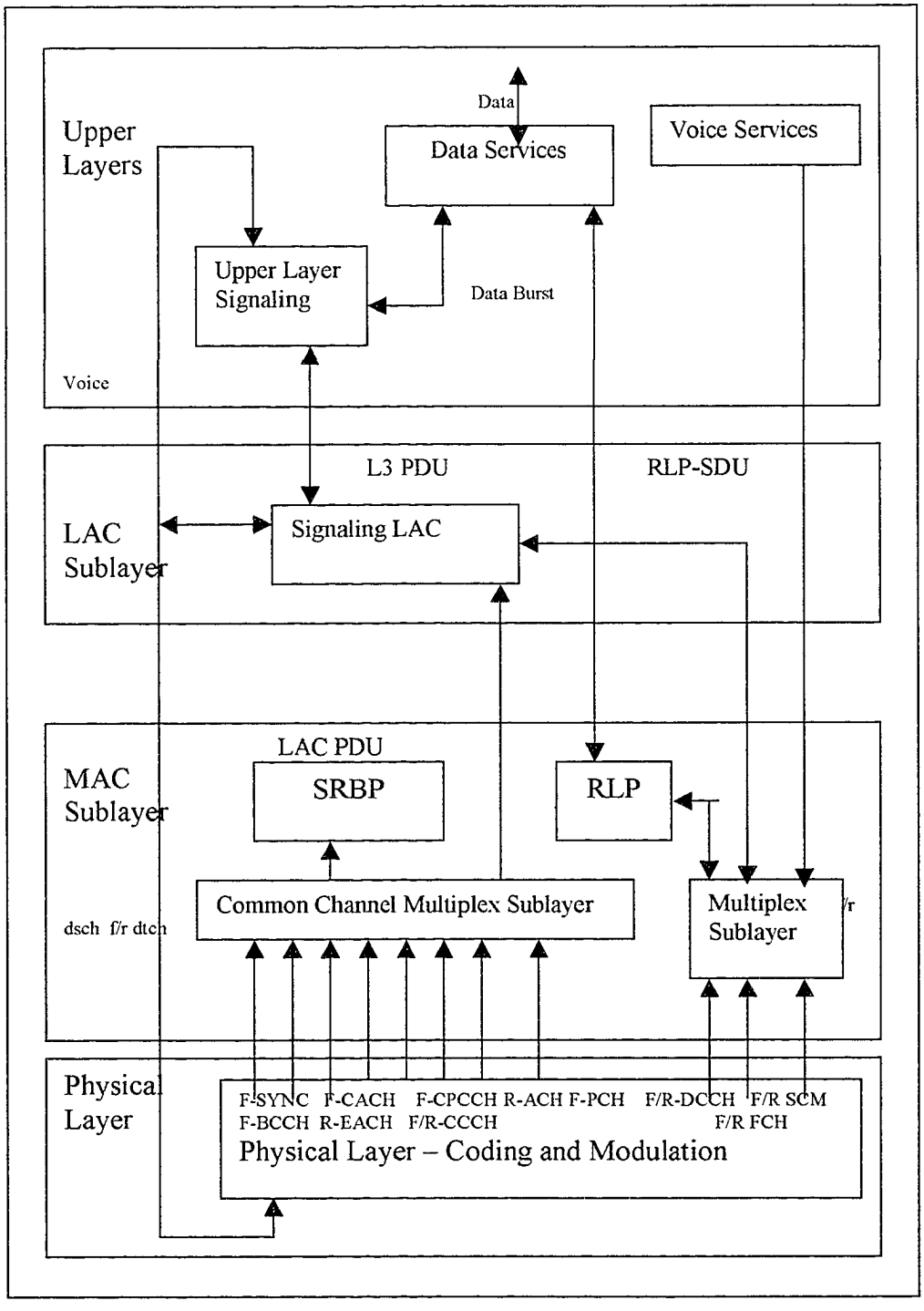


Figure 5.2. CDMA 2000 Architecture

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bibliography

5.1. "IMT-2000 (G3G) Strategy and Key Technical Issues". The Next-Generation Wireless Communications Conference. Far Eastone, Taiwan. 1999.

5.2. "Introduction to cdma2000 Standards for Spread Spectrum Systems". http://www.3gpp2.org/Public_html/specs/C.S0001-0_v1.0.pdf

5.3. Lin, Yi-Bing, Imrich Chlamtac. "Wireless and mobile Networks Architecture". John Willey and Sons, New York, NY. 2001

5.4. Nokia Website:. www.nokia.com/networks/17/gprs//gprs1.html, 2000. Harmonized Global 3G (G3G) Technical Framework for ITU IMT-2000 CDMA Proposal. May 1999.

5.5. Ojanpera, Tero, R. Prasad. "Wideband CDMA for 3rd Generation Mobile Communications". Artech House. 1998.

5.6 Trillium Digital Systems. General Packet Radio Service (GPRS) White Paper. May 1999.

5.7. UMTS official website. http://www.umtsworld.com/technology/fdd

5.8. Zvonar, Z., P. Jung. "GSM Communications Towards Third Generation Systems". Kluwer. Berlin, Germany. 1998

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 6. Beyond 3rd Generation Wireless Networks

There are already several papers trying to define "4th generation" wireless communication networks also referred as "Next Generation" networks.

6.1 Next Generation Wireless System architecture requirements and services

Next generations systems will inherit a non-homogenous infrastructure architecture, consisting of different switching fabrics and a large variety of physical media.

The next generation wireless network will have the following essential characteristics:

- Fully digital
- The backbone will be connectionless packet switching
- The variety of air interfaces inherited from 2nd and 3rd generation systems will have to accommodate packet switching technology
- Air interface will work at a speed of up to 100Mbps in different bandwith in the 5GHz, 17GHz and 60GHz range
- An overlaid architecture will provide seamless and transparent internetworking in the variety of air interfaces
- Public Wireless access will be possible at data rates of 20Mbps in dense urban areas, in offices and homes
- Service operators will provide non-real time wideband coverage, at 10Mbps range, in most public spaces, public transportation, and along highways

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Rural area will be covered at a limited 1Mbps rate

- Seamless transition from private networks to publicly operated networks will be possible

- **Ad-hoc networking will compete with traditional public operators**

- Technologies for fair and efficient multi-operator sharing of unlicensed spectrum will be in place

- The control of the new emerging ad-hoc networks, including routing and mobility, will be fully distributed and highly reliable

- Base stations/Access Ports in the public network will work in multi-mode, with multiple access air interfaces, in order to accommodate a large variety of mobile of fixed wireless terminals, and will be inexpensive

- Adaptive antenna arrays with self-configuring capabilities will be in use

- A large variety of terminals will be in use, from single-purpose cheap terminals designed for a specific service or function, to multi-service terminals with several functionalities and working in several standards

- Multi-service terminals will have to implement software-configurable hardware architectures

- The implementation of horizontal handover (between cells of networks with the same radio characteristics) and of vertical handover are both essential, in order to provide seamless connectivity for multi-service mobile terminals; The design of handover algorithms will have to include Quality of Service (QoS) criteria, as for example the availability of a higher bit rate bearer in a new cell.

- Terminals will operate in a speed range, from slow 10kbps to up to 100Mbps

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Battery capacity/weight/volume ratios will be up by one order of magnitude compared to those in use today

- Terminals operating in the 5GHz and 60GHz range will use advanced adaptive antennas

- New efficient protocols will be designed, tailored to the new wireless environment and able to conserve frequency spectrum and computing power in mobile terminals

- Seamless mobile multicasting will be an integral part of the network architecture

- The main services offered by next generation wireless networks will be:

- Multiparty video communication will be the major service for this generation. Real-time communication and multicast has to be adapted to the requirements of group video communication.

- Information provisioning, mainly dominated by educational or recreational material

- All car, office and home equipment and appliances will have wireless interface for remote control or remote diagnostics

- more sophisticated and reliable security in accessing already existing electronic services, as telebanking or telecommerce.

6.2 Next Generation Wireless Network Regulatory Issues

- The main issue is the level of standardization: the new systems should be fully standardized, just able to inter-work or good enough to ensure peaceful coexistence.

- An other important issue to decide is if the whole spectrum of bandwith will be all the time available or, as demand might only occur for short periods, should make spectrum bandwidth available as and when necessary.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- An other important issue to decide is if the whole spectrum of bandwith will be all the time available or, as demand might only occur for short periods, should make spectrum bandwidth available as and when necessary.
- Mechanisms for dynamic spectrum management need to be designed so that the radio spectrum can be switched from one type of service to another, as required.
- New policies need to be designed for accounting and billing of services, provided over a variety of access networks. Such policies will be simplified and tariffs will be rather fixed subscription based, than today's duration-based.

The Next Generation Wireless Network design is based on two complementary approaches: one based on evolution through a network-centric view and the other based on the user-centric view. AN ubiquitous connectivity to the Internet will be available not only for person to person communication, but as well with person to machine and machine to machine communication

In implementing next generation services, as first steps should be interworking between access networks and implementing enhanced versions of current technologies for broadcast, cellular and short range communications.

A second step will include the implementation of access technologies for transmission at more than 50 Mbit/s, for fast moving users as, well as ultra wide band systems for wide area coverage. Consequently, several interesting technological challenges and regulatory issues need to be addressed before the Next Generation Wireless Network vision becomes a reality.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bibliography

6.1 AT&T Labs. "4G Wireless Streaming Systems and Services"

http://www.research.att.com/news/2001/September/4G.html

6.2. http://www.s3.kth.se/radio/4GW/

6.3. http://www.s3.kth.se/radio/4GW/public/papers.html

6.4 Evans, B.G. Baughan, K. "Visions of 4G". http://www.ist-wsi.org/B_Evans.pdf.

6.5. http://www.fcc.gov/wrc03/files/docs/meeting/iwg/iwg_1/

wrc03_iwg_1_background_doc_13.pdf.

6.6. IEEE, Mohr, Werner. Becher, Reinhard. "Mobile Communications

Beyond Third Generation".

6.7. "IP for 3G". Wiley and Sons. 2003.

6.8 Karlson, B. "Wireless Foresight". Wiley and Sons. September 2003.

6.9. Mohr, W., W. Konhauser. "Access Network Evolution Beyond 3rd Generation

Mobile Communications". IEEE communications Magazine. 2000.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 7. Wireless Intelligent Networks (WIN)

7.1. Intelligent Networks

For almost a century, since sir Alexander Graham Bell invented the telephone, the telecommunication system evolved slowly around POTS (Plain Old Telephone System). The invention and evolution of electronic computing systems had a remarkable influence in the evolution of telecommunication system, which started acquiring intelligence.

One can notice 3 major milestones toward the radical change into the today Intelligent Telecommunication Networks:

1. In 1965 the introduction of "Stored Program Control" , and consequently the computer controlled switching systems. The telecommunication network becomes digital in two aspects:

- digital control of voice and data channels
- digital representation of if the information

2. In 1976 the implementation of the Common Channel Interoffice Signaling System: the separation of control signaling for many channels in a separate channel. By this, "outband" signaling becomes digital, standardized and secure.

3. In 1984 it was implemented Service Independence: switching systems are provided with "primitives" (small service independent program utilities) used as building blocks to create new services (ex. authentication, interaction, screening, number translation, charge etc.).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The traditional network functions are complemented by a set of specialized functions, which enable the introduction and marketing of new services.

Intelligent Network Architecture

Several IN architectures, common in concepts but different in terminology, have been defined, or are under development.

The pioneering work on INs has been done by Bellcore (today Telcordia Technologies), with the proposition of Intelligent Network-1 (IN-1), with a first implementation at Bell South, in 1987.

IN-1 was initially designed for the provision of "800 Service", and enlarged with a few other services.

One year later, in 1988, new architectural features were added and intermediate stage toward full service independence was created under the name of IN-1+.

Later IN-2 was conceived as an Open Network Architecture, with full service independence, implying radical changes in the service switching nodes.

In 1989 a new IN architecture named Advanced Intelligent Network, was proposed, and a replacement of IN-1 and IN-2, since a smooth evolution with the existing network equipment was not possible.

In parallel with Bellcore's initiative the International Telecommunications Union (ITU, formerly CCITT) in Study Groups 9 and 18, started the development of the Intelligent Network Long Term Architecture (IN-LTA), conceived as a series of upwardly compatible IN Capabilities Set (CS); each capability set is defined in terms of the services to be supported and the functional architecture supporting these services.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Intelligent Network is more than just a new network architecture, introducing a service-oriented view: it is a complete framework for the uniform creation and provision of telecommunication services, and an important step ahead in the unification of telecommunication and computer networking.

Intelligent Network Services

Over 70 AIN services are already in use and over 280 are designed for near-future implementation.

Between the most exciting data and video services we can mention:

- viewer dependent access to content
- viewer customization of available channels
- pay-per-view subscription on-demand
- home bound shopping
- inbound/outbound message screening
- multiple internet protocol addresses
- uniform session distribution
- icon driven access to material
- television-based internet access
- TV/Voice/Internet correlation.

Definition of Intelligent Network

We can define IN as a service oriented architectural concept with a large level of generality, providing creation, management and real-time execution of a variety of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

network services, in a distributed environment, creating 3 degrees of freedom: service, network and vendor independence.

Intelligent Networks create and support a multilayer communications environment, in which a variety of services will be provided by multiple service providers in a competitive way.

Intelligent Network Conceptual Model

The international standardization of IN started in 1989:

- in ITU study Groups 9 and 18 (ITU Recommendations Q.1200)
- as well as in European Telecommunications Standardization Institute (ETSI) Group NA6 (ETSI TCR-TR NA-60106).

The reference model is known as Intelligent Network Conceptual Model (INCM), describing the capabilities and characteristics of an IN-structured network.

The basic idea is to define a top-down approach for the definition of IN, similar to the OSI model for network protocols.

Four structures are identified (see Figure 7.1):

- The upper two planes of the INCM focus on service creation and implementation by using generic building blocks, providing service independence.
- The lower two planes define a generic IN service provisioning architecture, providing thus network independence.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

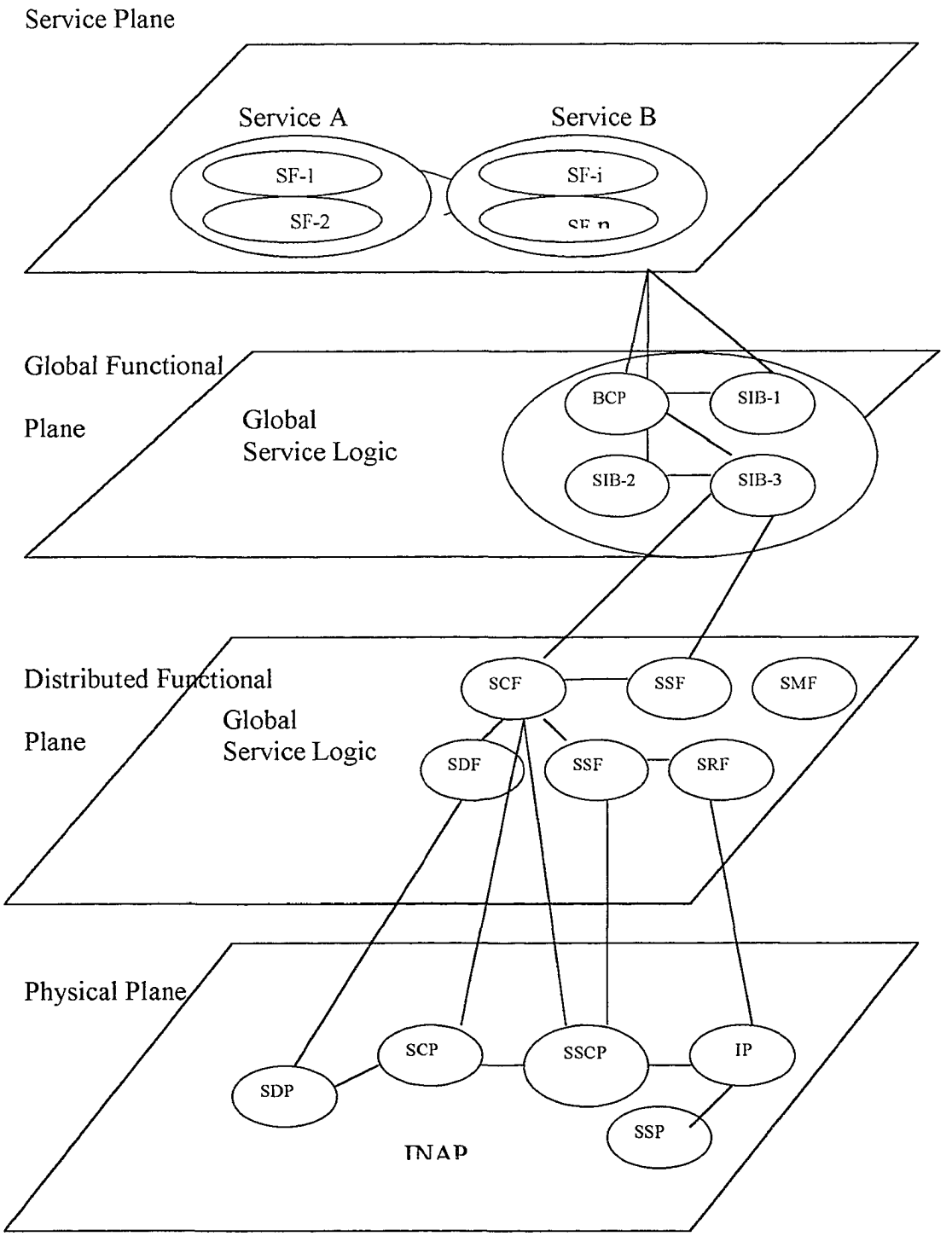


Figure 7.1. Intelligent Network General Architecture

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. Service Plane (SP):

- describes services from a user's perspective

- each service consist of one or more generic blocks named Service Features (SF)

2. Global Functional Plane (GFP):

- models the network as a single programmable entity

- contains the Service Independent Building-blocks(SIB), used as reusable network capabilities, for efficient service feature implementation

- Basic Call Process (BCP) represents interactions between dedicated SIB

- Global Service Logic (GSL) describes how SIBs are chained together using BCP, for a Service Feature (SF).

3. Distributed Functional Plane (DFP):

- models the distributed view of an IN in term of units of Network Functionality; NF is a set of Functional Entities (FE) as:

- Service Switching Function (SSF)

- Service Control Function (SCF)

- Service Data Function (SDF)

- Service Management Function (SMF)

- Specialized Resource Function (SRF)

- each Functional Entity (FE) may perform a variety of Functional Entity Actions (FEA), id est each SIB is implemented by a particular sequence of FEAs, performed by specific FEs.

- some of FEAs result in Information Flow (IF), I.e. message exchanges in case where different functional entities cooperate for the provision of SIBs.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- provides transparency to the Physical Network

4. Physical Plane (PP):

- models the physical aspects of an IN-structured network
- defines different Physical Entities(PE) and their interfaces (as SSP, SCP, SDP,IP), where the corresponding FEs from DFP are located; one single Functional Entity (FE) must be implemented with one single Physical Entity (PI), but different FEs can be implemented with the same PI
- The Information Flows (IF) in the DFP are implemented in the PP through a standard protocol: the IN Application Protocol (INAP), i.e. Physical Entities (PE) communicate by means of INAP for IN service execution.

IN architecture definition

Using this model, in order to define an IN architecture, the following steps have to be followed:

- Definition of Services, by identifying the Service Features which can be combined in order to build up each service
- Decomposition of Service Features in Service Independent Building-blocks (SIB)
- Definition of the IN architecture, in two stages:
  - Network Entities are defined in terms of functional elements and their interactions
  - Network Entities are allocated to specific physical entities.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.2. Advanced Intelligent Networks (AIN)

AIN was a parallel IN standardization activity in North America, driven by Bellcore (today Telcordia) and implemented by Regional Bell Operating Companies (RBOCs) in the 1992-1998 time frame. AIN has two implementation stages of Release 1: Release 0.1 and Release 0.2.

AIN Components

Six basic components of AIN were conceived:

1. Service Switching Points (SSP):
   - permit to detect AIN activation events from users
   - communicate with other AIN elements
   - control resources necessary to process AIN calls
2. Service Control Points (SCP):
   - contains the application programs which provide implemented AIN services
   - responds to messages from a switch, with AIN specific switching capabilities, in order to invoke these application programs
3. Adjunct Processor (AP):
   - is a mate of SCP and is functionally equivalent; the difference consist in deployment strategy: the Adjunct is connected to AIN switches via high speed data links, rather than CCS/SS7 links, making it more suitable to support services which require short response time
4. Signal Transfer Points (STP):
   - **controls the link to SSP, SCP or IP through CCS/SS7 network**

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5. Intelligent Peripherals (IP):

- provide AIN service capabilities that communicates with the SSP via the IP interface.
- can be used to provide resources such as the collection of digits, play announcements, provide stand-alone services, and originate/terminate calls

6. Service Management System (SMS):

SMS is an off-line support facility used to enter:

- customer data into SCP data base
- call-processing software in SCP
- communicate with SCP through interface cards using BX.25 protocol.

Basic Call Model (BCM)

Basic Call Model is a representation of the processing flow within an SSP for a basic two party call.

BCM has two parts:

1.the Originating BCM and

2. the Terminating BCM.

- The Originating BCM deals with the processing related to the originating (calling) party.
- The Terminating BCM deals with the processing associated with the called party.
- Trigger Detection Points(TDP) are associated with the BCM points in call processing. This enables AIN services to be invoked via Service Logic (SL) at the SCP/Adjunct, when certain trigger criteria are met.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Service Creation

The AIN architecture is built on the premises that multiple developers will be able to build and modify services in a flexible and responsive way, making service creation independent from service delivery.

Service creation function must be defined in such a way as to insure:

- network reliability, integrity and privacy;
- protection from feature interactions;
- protection from a variety of competing service applications.

Independent administrative groups will monitor and manage new service implementation.

Future Directions

Some of the most important Advanced Intelligent System architectures and features already envisioned, are:

- global AIN: to fulfill social needs and economic viability
- network based educational systems
- integrated medical systems
- multifunctional intelligent networks (integration of computation with communication)
- home intelligent networks (PC based)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.3. Wireless Intelligent Networks (WIN)

Wireless intelligent network (WIN) is an implementation of the concepts and services of Intelligent and Advanced Intelligent Networks in the area of Wireless Networks. The main concepts of WIN are being developed by the Telecommunications Industry Association (TIA) Standards Committee TR45.2. The purpose of this committee is to implement intelligent network (IN) capabilities, based on interim standard (IS)-41, into wireless networks. IS–41 (presented above) facilitates roaming. Basing WIN standards on this protocol, enables a graceful evolution to an IN without making current network infrastructure obsolete.

WIN overview

After a successful mass deployment of wireless communications, customers are requesting services beyond the completion of a simple call. Subscribers demand innovative ways to use the wireless phone. They want multiple services that allow them to handle or select incoming calls in a variety of ways.

New services are conceived every day; enhanced services like: caller ID, prepaid service, voice/data messaging, internet surfing, or location-sensitive billing, will make a difference in a competitive, service-provider market.

Initially, Intelligent Network (IN) solutions have revolutionized wireline networks. Wireless Intelligent Network (WIN) will bring these strategies into the wireless networks. The evolution of wireless networks to a WIN concept of service deployment delivers the following advantages, similar to the wireline IN benefits:

- multi-vendor product offerings that foster competition

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- uniform services to subscribers across service areas

- efficient network utilization

- rapid service creation and deployment

Service Requirements particular to WIN

Mobility dictates a need for technology or standards that make it possible for different networks to talk to each other. The most important WIN services envisioned, are:

1. Roaming

Roaming is one of the factors driving the WIN standards to be developed and implemented.

Customers can roam out of their local calling area or out of their service provider's area. Both of these situations require messaging data before the call is even put through to handle the setup, authorization, and proper billing for roaming services. This service require signaling system 7 (SS7) messages, specific to IN networks, to be exchanged between the different devices of the system. Intelligent networking is required for various validations and billing reciprocation of wireless calls.

2. Carrier Select

Wireless providers are making agreements with one another, because of business partnerships and because carriers have licenses in different area.

Carrier-select services allow providers to select, based on codes or automatic handset selection, the network that will be used to complete the call. This way, a carrier can secure calls with its business partners.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

These services also allow subscribers the choice of programming new, sophisticated handsets to route calls selectively, which could save them money if they know different network rates.

3. Hands-Free Operation

Hands-free wireless services are extremely important for the safety of the mobile user. Related features like voice-activated dialing and feature activation require special technology that converts voice into data (speech to text and text to speech conversion). The call is routed to the Intelligent Peripheral (IP) that provide the special technology, such as voice recognition, necessary for hands-free, voice-controlled services. To get the message or voice signals routed to the devices that collect that information and translate it to data information requires special routing specific to intelligent networking. An IN application sends the message to the device, turns it on, translates the digits, removes it from the telephone circuit, and lets the call go on. As a result, an IN application is an efficient utilization of IP facilities.

4. Fee Structure

Calls are being handed off among networks owned by different network operators and carriers. After the call is handled properly, billing takes place. WIN flags can be written right into the call record so that billing reflects the specific call handling. With so many different agreements, carriers may have negotiated different fee structures with each partner. Using an WIN flag in the billing record makes processing fees between carriers much easier.

Wireless companies have always paid wireline companies for handling calls that terminate outside the wireless network—a majority of the calls. Money has always gone

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

one way. As a result of the Telecommunications Act of 1996, billing can go both ways. Wireless companies can now get paid for calls coming into the wireless network. These changes make billing relationships more complex, and will increase the need for WIN flags.

Providers also will offer services such as calling party pays to make it easier for a wireless subscriber to receive calls at no charge. Again, these types of services drive a need for more centralized WIN capabilities.

5. Data-Service Capabilities

Handset displays allow customers to use various messaging services. One, called short message service (SMS), works much like a pager. It allows phones to send and receive messages in addition to making or taking telephone calls. SMS requires many SS7 messages just to set up the signaling and the mechanism to get the data through the wireless network. It requires a significant amount of checks and balances, finding the database, pulling up the message, encapsulating it with the right header information to route it to the correct user, and finally sending it out like a phone call. It may even go out over a data channel so that the user can receive a message while making a voice call at the same time. This gets complex and requires IN routing and authentication.

6. Voice-Controlled Services

Voice-controlled services employ voice-recognition technology to allow the wireless user to control features and services using spoken commands, names, and numbers. There are two main types of automatic speech recognition (ASR). Speaker-dependent requires specific spoken phrases unique to an individual user. Each user is required to train the ASR system by recording samples of each specific phrase. The other is speaker-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

independent ASR, which requires the use of specific spoken phrases that are independent of the speaker. The individual user need not train the system.

6.1 Voice-Controlled Dialing (VCD)

VCD allows a subscriber to originate calls by dialing digits using spoken commands instead of the keypad. VCD may be used during call origination or during the call itself.

6.2 Voice-Controlled Feature Control (VCFC)

VCFC permits a calling party to call a special VCFC directory number, identify the calling party as an authorized subscriber with a mobile directory number and personal identification number (PIN), and specify feature operations via one or more feature-control strings. This service is similar to remote feature control (RFC) except that the subscriber is allowed to dial feature-control digits or commands using spoken words and phrases instead of keypad digits.

6.3 Voice-Based User Identification (VUI)

VUI permits a subscriber to place restrictions on access to services by using VUI to validate the identity of the speaker. VUI employs a form of ASR technology to validate the identity of the speaker rather than determine what was said by the speaker. VUI requires that the subscriber register the service by training the ASR system by recording a word or phrase. When a user attempts to access a service, the ASR system prompts the user to say the special phrase.

7. Incoming Call-Restriction/Control

Incoming calls to a subscriber may be given one of the following termination treatments: the call is terminated normally to the subscriber with normal or distinctive alerting; it is forwarded to voice mail or to another number; it is routed to a subscriber-specific

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

announcement; or it is blocked. These kinds of services help subscribers control incoming calls and their monthly airtime bills. From a marketing standpoint, they entice cost-conscious customers who might not want unlimited access from callers.

8. Calling Name Presentation (CNAP)

CNAP provides the name identification of the calling party (e.g., personal name, company name, restricted, not available) to the called subscriber. The calling name information (CNA) is derived from the calling number information (CNI), which is generally provided to the terminating network as part of the basic call setup. Optionally, the date and time of the call may be provided to the called subscriber.

9. Password Call Acceptance (PCA)

PCA is a call-screening feature that allows a subscriber to limit incoming calls to only those calling parties who are able to provide a valid password (a series of digits). Calls from parties who cannot provide a valid password will be given call refusal while PCA is active.

10. Selective Call Acceptance (SCA)

SCA is a call-screening service that allows a subscriber to receive incoming calls only from parties whose calling party numbers (CPNs) are in an SCA screening list. Calls without a CPN will be given call-refusal treatment while SCA is active.

11. Data Capability

The main data capabilities are:

11.1 Short Message Service (SMS)

SMS provides the ability to deliver short messages as a packet of data between two service users, known as short message entities (SMEs). SMS incorporated into PCS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

networks allows for simultaneous paging and voice. Among its applications are paging via wireless phone screens and voice-mail notification.

11.2 Speech-to-Text Conversion (STC)

STC permits a calling party to create a short alphanumeric message by speaking to an ASR device that will perform speech-to-text conversion. The short message may then be distributed by any means available such as short message delivery.

12. Prepaid Services

Prepaid cellular services can take a number of forms, like:

- debit card

- connection to a smart card.

These services allow customers to pay before they call and not be billed later. As the subscriber has already paid for the service, the carrier is not burdened with the risk or overhead of payment collection.

Functional Components of WIN

The WIN is an specific implementation of the wireline IN mode. But, the distinction between the wireless and wireline network is that many of the wireless call activities are associated with mobility, not just the actual phone call. In the WIN, more call-associated pieces of information are communicated between the MSC and the SCP or HLR. The WIN moves service control away from the MSC and up to a higher element in the network, usually the SCP (Figure 1).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

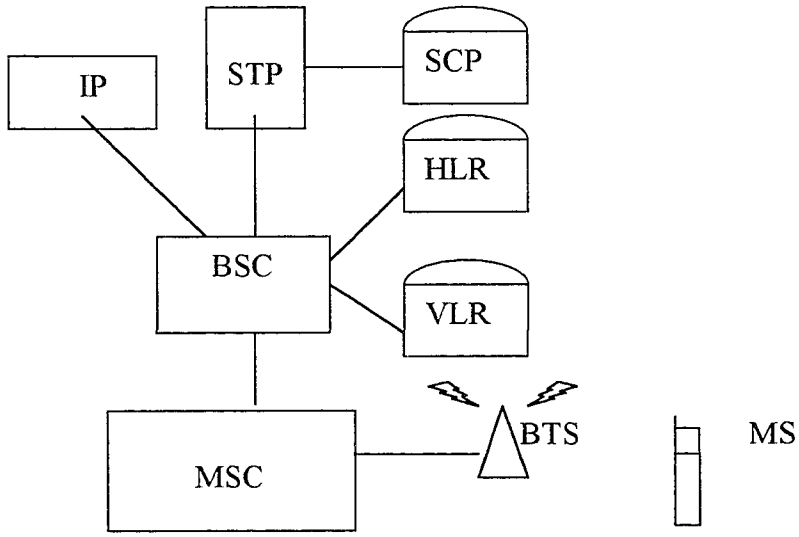


Figure 7.2. Wireless Intelligent Network (WIN) Basic Architecture

The main components of the WIN are:

1. MSC as Service Switching Point (SSP)—In the IN, the SSP is the switching function portion of the network. The mobile switching center (MSC) provides this function in the WIN.

2. Service Control Point (SCP)—This device provides a centralized element in the network that controls service delivery to subscribers. High-level services can be moved away from the MSC and controlled at this higher level in the network. It is cost-effective because the MSC becomes more efficient, does not waste cycles processing new services, and simplifies new service development.

3. Intelligent Peripheral (IP)—The IP gets information directly from the subscriber, be it credit-card information, a PIN, or voice-activated information. The peripheral gets information, translates it to data, and hands it off to another element in the network—like the SCP—for analysis and control.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4. Signal Transfer Point (STP)—This is a packet switch in the signaling network that handles distribution of control signals between different elements in the network such as MSCs and HLRs or MSCs and SCPs. The advantage of an STP is that it concentrates link traffic for the network. It can also provide advanced address capabilities such as global title translation and gateway screening.

5. Location Registers—These are used to supplement MSCs with information about the subscriber. The number of subscribers that the switch supports changes as roamers move in and subscribers move to other switches. The database of active subscribers changes very dynamically. Each MSC cannot have the database for all potential users of that switch. The following location registers help to get around that problem:

5.1 Visitor Location Register (VLR)—Within an MSC there is a VLR that maintains the subscriber information for visitors or roamers to that MSC. Every MSC or group of MSCs will have a VLR.

5.2 Home Location Register (HLR)—Information on roamers is obtained from that subscriber's HLR. Each subscriber is associated with a single HLR, which retains the subscriber's record. When the subscriber roams to another switch, the VLR queries the subscriber's home HLR to get information about that subscriber. When a phone call goes to a subscriber's MSC, the MSC recognizes that the subscriber is roaming and asks the HLR for the subscriber's location. The HLR will communicate that information to the VLR and relay a temporary location number received from the visited system. In the WIN architecture, the HLR is usually a network element such as an SCP.

WIN Call Model (W-CM)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The WIN call model enables the network to handle new triggers (which are decision points in a call) and new transaction capability application part (TCAP) messages.

The evolution to WIN will be a major step forward for North American wireless networks. It involves an industry-wide consensus among equipment vendors and service providers to incorporate IN concepts into existing wireless networks.

The following steps will need to occur before WIN will be a reality:

- incorporation of SCP, IP, and SN into the wireless network architecture
- evolution of the MSC to a SSP
- separation of call control and transport from service control
- development of generic call models, events, and trigger points

Case1. A Stand-Alone HLR represent the first step in a WIN implementation strategy. As subscriber bases grow and technology changes, networks can be expanded easily with a stand-alone HLR. When a provider migrates to a stand-alone HLR, from the MSC to an independent IN network node, the MSC can be dedicated to efficient call processing. In such a way the complexity of the network is significantly reduced.

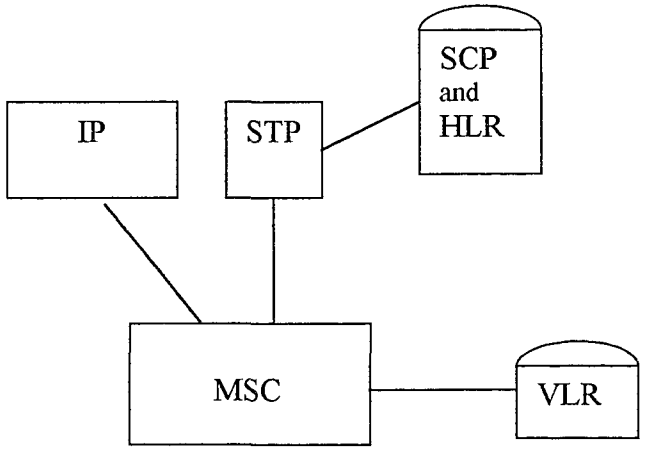


Figure 7.3. WIN with a Stand-Alone HLR

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

There are two reasons for making the HLR a network element:

MSCs are switches containing a certain amount of processing power. An HLR performs computation, not switching. When the HLR is taken out of the MSC to be an independent network node, MSC has more time for its main functionality, to switch phone calls.

To provision an MSC with subscribers, it is necessary to put the data in each MSC, and each one could have a different. With a unified HLR platform, there is only one provisioning mechanism for the HLR database with a standard protocol.

Moving to a stand-alone HLR is the first step in the implementation of WIN services. When IN architecture is implemented, the network is automatically set up for interoperability with other services and other networks as the WIN standards come into play.

There are several other advantages of IN–based centralized service control, including the following:

- reduced delivery time for services
- reduced costs for service deployment
- flexibility that makes elements such as fraud control and specialized services more easily deployed
- reduced network component costs

Case2. SCP/HLR merge: As the HLR becomes a separate element in the network, it will resemble more to an SCP. Therefore, the SCP and HLR functions could merge, and an unified SCP/HLR in the network becomes the next logical step.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

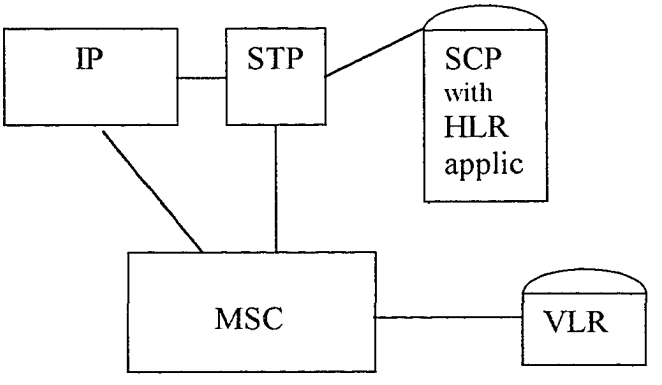


Figure 7.4. WIN with a merged SCP/HLR

7.4. Evolution of WIN Networks

The movement to develop a WIN standard was originated wireless network operators, through the Cellular Telecommunications Industry Association (CTIA). They developed a set of requirements that defined the new network architecture, incorporating the service flexibility of wireline Intelligent Networks (IN)s with the mobility aspects of wireless networks.

WIN Phase-1, the first phase of WIN standards, was published in 1999 and established the fundamental call models and operations required to support this flexible service architecture. Many service providers have implemented WIN Phase-1 in their networks. Examples of WIN Phase 1 services are calling name presentation and restriction, call screening, or voice-control services.

WIN Phase-2 standards provide both additional service capabilities for wireless operators as well as greater harmonization of network capabilities and operations with emerging third-generation network requirements. WIN Phase 2 includes:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MSC triggers for an IN prepaid solution or enhancements to support location-based services.

These requirements are based on four service drivers:

- location-based charging,

- fleet and asset management service,

- enhanced call routing service, and

- location-based information service.

Bibliography

7.1. Ahamed, Syed V., Victor B. Lawrence. "Intelligent Broadband Multimedia Network". KluwerAcademic Publishers. Boston, USA. 1997

7.2. Ahamed, Syed V., Victor B. Lawrence. "Design and Engineering of Intelligent Communications Systems". Kluwer Academic Publishers. Boston USA. 1997.

7.3. Bellcore Special Report: "Advanced Intelligent Network release 1, Network and Operations Plan". SR-NPL-001623, Issue 1, June 1990.

7.4. Black, Uyless. "The Intelligent Network Customizing Telecommunication Networks and Services". Prentice Hall, Upper Saddle River, NJ. 1998.

7.5. Christensen, Gerry, P.G. Florack, R. Duncan, "Wireless Intelligent Networking", Artech House. 2001.

7.6. Faynberg, Igor, L. R. Gabuzda, M. P. Kaplan, N. J. Shah. "The Intelligent Networks Standards and Their Application to Services". McGraw-Hill. 1997.

7.7. Garrahan, J.J.. "Intelligent Network Overview", IEEE Communications Magazine 31(3), pg. 30-37. March 1993

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.8. IEEE Special Issue: "Marching Toward the Global Intelligent Network". IEEE Communications Magazine, 31(3). March 1993.

7.9. Lobley N.C. "Intelligent mobile networks". BT Technologies, J Volume 13, No2, April 1995.

7.10. Magendanz, Thomas, R. Popescu-Zeletin. "Intelligent Networks: Basic Technology, Standards and Evolution". Thomson Computer Press. 1996.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 8. Wireless Data Communication Networks

8.1 Classification

The major types of wireless data communications networks include the following:

- Wireless Local Area Networks (LANs)
  - 802.11b, sometimes called wireless Ethernet
  - 802.11a, HiPerLAN II, and HomeRF, seen as alternatives to 802.11b
- Fixed Broadband Wireless Multiservice Wide Area Networks (WANs)
  - MMDS and LMDS
  - Small Dish Satellite
- Mobile Wireless (voice/telephony with increasing integration of data and video)
- Wireless Personal Area Networks (mainly based on Bluetooth™ technologies).

There are already 3 main standards in place to assure compatibility of wireless communications in in-door applications: IEEE 802.11, HyperPlan and IrDA.

New emerging standards are complementing the existing ones: Bluetooth, Home_RF and Wireless_ATM.

Standards

8.2. Wireless ATM (WATM)

WATM is a technology developing as a result of a close cooperation between the ATM forum and the ETSI broadband radio access forum (BRAN). The aim is to develop 3 network standards:

- HIPERPLAN-2

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- HIPERACCESS
- HIPERLINK.

All are based on ATM technology, and are capable to support broadband services in a mobile environment.

8.3. Infra-red Data Access (IrDA)

The aim of Infra red Data Access network is to ensure serial data-communications through free-space Infra-Red channels. It is intended to be used for point-to-point communication, at the view-of-sight, between a large variety of devices.

In 1994 IrDa approved IrDa standard 1.0 which includes:

- Serial Infra-Red data-link standard (SIR)
- Infra-red Link Access Protocol (IrLAP)

It was intended to support ad-hoc or walk-up connectivity applications. It support speeds in the range of 2.4kbs to 115Kbs.

In October 1995, version 1.1 included access speeds of up to 4Mbs.

IrDa complying systems are made up of 4 elements:

- IrDA enabling application
- IrDA software protocols stack
- IR controller and
- IR transceiver.

If physical and data-link layer level are designed in detail, upper layers have few design restrictions, for greater flexibility.

IrDa structure can be divided as follows:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. Physical level, manages the IR transmitter and receiver, and performs error detection. Also, the IrDA Dongle Interface presents the specifications of connectors and pin assignments for interoperation with desktop systems.

2. IrLAP is dealing with data-link level communication with similar devices. It implements:

- the station discovery process and address assignment.
- controls the QoS negotiation procedures.

It supports 2 types of communications modes:

- connectionless oriented: implementing a broadcast transmission and
- connection oriented: based on a token based system.

As an alternative for simpler applications at this level, the IrLITE protocol was conceived. It presents a IrDa link with simpler communications needs.

3. IrLMP optimizes the IR link providing the possibility of several logic channels that share the same IrLAP link service. Also introduces an information service that allows one device to see the services available in other devices.

8.4. IEEE 802.11 Standard for Wireless LAN

The Institute of Electrical and Electronics Engineers (IEEE) ratified the original 802.11 specification in 1997 as the standard for WLANs. That version of 802.11 provided for 1 Mbit/s and 2 Mbit/s data rates and a set of fundamental signaling methods and other services. The data rates supported by the original 802.11 standard were too slow to support most general business requirements with and did little to encourage the adoption of WLANs. Recognizing the critical need to support higher data-transmission rates, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

fall of 1999 saw the IEEE ratify the 802.11b standard (also known as 802.11 High Rate) for transmissions of up to 11 Mbit/s.

802.11 defines two types of equipment, a wireless "station", which is usually a PC equipped with a wireless network interface card (NIC), and an "access point" (AP), which acts as a bridge between the wireless and wired networks. An access point usually consists of a radio, a wired network interface (such as IEEE 802.3), and bridging software conforming to the 802.1d bridging standard.

- **The access point** acts as the base station for the wireless network, aggregating access for multiple wireless stations onto the wired network.

- Wireless end stations can be 802.11 PC Card, PCI, or ISA NICs, or embedded solutions in non-PC clients (such as an 802.11-based telephone handset).

The 802.11 standard define two working modes: "infrastructure" mode and "ad hoc" mode.

- **In infrastructure mode**, the wireless network consists of at least one access point connected to the wired network infrastructure and a set of wireless end stations. This configuration is called a Basic Service Set (BSS). An Extended Service Set (ESS) is a set of two or more BSSs forming a single subnetwork. Since most corporate WLANs require access to the wired LAN for services (file servers, printers, Internet links) they will operate in infrastructure mode.

- **Ad hoc mode** (also called peer-to-peer mode or an Independent Basic Service Set, or IBSS) is simply a set of 802.11 wireless stations that communicate directly with one another without using an access point or any connection to a wired network. This mode is useful for quickly and easily setting up a wireless network anywhere that a wireless

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

infrastructure does not exist or is not required for services, such as a hotel room, convention centre, or airport, or where access to the wired network is barred (such as for consultants at a client site).

The three physical layers were originally defined in 802.11: two spread-spectrum radio techniques and a diffuse infrared specification. The radio-based standards operate within the 2.4 GHz ISM band. These frequency bands are recognized by international regulatory agencies, such as the FCC (USA), ETSI (Europe), and the MKK (Japan) for unlicensed radio operations. For this reason, 802.11-based products do not require user licensing or special training. Spread-spectrum techniques, in addition to satisfying regulatory requirements, increase reliability, boost throughput, and allow many unrelated products to share the spectrum without explicit co-operation and with minimal interference.

The original 802.11 wireless standard defines data rates of 1 Mbit/s and 2 Mbit/s via radio waves using two different - mutually incompatible - spread spectrum transmission methods for the physical layer:

With Frequency Hopping Spread Spectrum (FHSS), a transmitting and a receiving station are synchronized to hop from channel to channel in a predetermined pseudorandom sequence. The prearranged hop sequence is known only to the transmitting and receiving station. In the U.S. and Europe, IEEE 802.11 specifies 79 channels and 78 different hop sequences. If one channel is jammed or noisy, the data is simply retransmitted when the transceiver hops to a clear channel.

Under Direct Sequence Spread Spectrum (DSSS), each bit to be transmitted is encoded with a redundant pattern called a chip, and the encoded bits are spread across the entire available frequency band. The chipping code used in a transmission is known only to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sending and receiving stations, making it difficult for an intruder to intercept and decipher wireless data encoded in this manner. The redundant pattern also makes it possible to recover data without retransmitting it if one or more bits are damaged or lost during transmission. DSSS is used in 802.11b networks.

At the OSI data link layer 802.11 uses a slightly modified version of the CSMA/CD protocol known as Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) or the Distributed Co-ordination Function (DCF). CSMA/CA attempts to avoid collisions by using explicit packet acknowledgement (ACK), which means an ACK packet is sent by the receiving station to confirm that the data packet arrived intact.

It works as follows: A station wishing to transmit senses the air, and, if no activity is detected, the station waits an additional, randomly selected period of time and then transmits if the medium is still free. If the packet is received intact, the receiving station issues an ACK frame that, once successfully received by the sender, completes the process. If the ACK frame is not detected by the sending station, either because the original data packet was not received intact or the ACK was not received intact, a collision is assumed to have occurred and the data packet is transmitted again after waiting another random amount of time.

CSMA/CA thus provides a way of sharing access over the air. This explicit ACK mechanism also handles interference and other radio-related problems very effectively. However, it does add some overhead to 802.11 that 802.3 does not have, so that an 802.11 LAN will always have slower performance than an equivalent Ethernet LAN.

The standard 802.11 is in fact a wireless implementation of the wireline Ethernet protocol.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

802.11 Layers

802.11 is structured in the following layers:

1. Physical layer:

The physical layer is subdivided in 2 sub-layers:

- Medium Dependent (PMD)
- PHY convergence (PLCP)

2. Data-link layer:

The data link layer is subdivided in :

- Medium Access Control (MAC)
- Logical Link Control (LLC)

The many features of this standard in order to establish a reliable communication, made it fairly complex, the implementation cost being higher than the wireline counterpart. Both structured networks and ad-hoc networks, asynchronous and time-critical applications are supported.

---

1. Physical Layer

There are 2 specifications for the physical layer of the 802.11 protocol:

- 1.1 Frequency Hopping Spread Spectrum and
- 1.2 Direct Sequencing Spread Spectrum.

Both are using spread spectrum technology and use the non-licensed band around 2.4GHz.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Unlicensed bands like ISM (Industrial, Scientific, Medical,) have larger flexibility in use, the main restriction being the emission power (usually below 1W).

Spread Spectrum Techniques

SS systems use for transmission more bandwith than needed. To achieve this, usually a code is used by the transmitter prior to modulation, independent of the transmitted data, and known by the receiver. SS are more difficult to detect, intercept and decode and less sensitive to interference.

The most widely used SS techniques are: Frequency Hopping (FH) and Direct Sequencing (DS).

1.1 Frequency Hopping Technique

FH systems use conventional modulation techniques, but the carrier frequency is changed at a given rate, following a given sequence. This sequence represents the code of the system. The receiver must know the code in order to de-code the sequence.

- If the Hopping Rate is faster than the bit rate, then the system is named FFH (Fast Frequency Hopping).

- If the Hopping Rate is slower than the bit rate, then the system is named SFH (Slow Frequency Hopping).

FFH is very complex, and therefore is seldom in commercial use.

In a SFH system the bit stream is split into packets, which are each transmitted in a burst, with a different carrier frequency. Within a given burst the transmission is narrowband, using just the bandwith needed according to the bit-rate and modulation format.

In a noisy environment, there are some advantages of this type of transmission, like:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in case of a narrowband interference, this affects only those bursts who's carrier coincides with the carrier of the other transmitter, which doesn't happen to often.

If two FH systems are working in the same area, interference will occur each time the bursts of both systems are the same in the same carrier.

In a multipath environment, when the signal arrives to the receiver from different paths, with different delays, then frequency-selective fading occurs. As a result, some bursts will suffer fading but some other will be reinforced and most of them will not be affected. The affected burst is recovered, by forward error coding or retransmissions.

This type of transmission is useful until the loss is small. A small loss of bursts can be achieved by fighting against interference: the higher the number of available frequencies, the smaller the number of lost bursts.

In order to limit selective fading it is also necessary to limit the probability to have many of them simultaneously affected by fades.

FH systems are not appropriate to use when the spectrum is heavily used.

When frequency use is more important, then orthogonal FH can be used. This a techniques that allows a group of transmitters to use the same frequency band, all of them with FH, but with the hop sequences tightly synchronized, as it's impossible for two transmitters to use the same frequency at the same time. (this technique is used for example in the GSM).

If N is the number of used frequencies, then the number of transmitted burst, lost due to interference is between 1/N and 2/N. If N is bigger the loss is smaller. But if N increases then bandwith also increases.

In FH the following modulation techniques are generally used: FM, FSK and QPSK.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Because the switching time between frequency IDs is a time lost for transmission, it has to be minimized. In addition, it has to be made much smaller then the length of the burst. Receivers have to synchronize their frequency hop sequence with the one embedded in the received signal. There are several synchronization techniques and all arte dependent on the hop rate and the modulation format.

1.2. Direct Sequence Technique

In DS systems, the modulation rate is intentionally increased. This is achieved by combining the bit sequence with a higher rate binary sequence (called chip sequence) to obtain a new sequence with the chip rate. This is then used to modulate the carrier.

At the receiver side: the signal is demodulated and then it is recombined with the same chip sequence to restore the original data.

Their implementations require to determine:

- what kind of chip sequences to be used
- how to combine them with the data sequence
- what modulation formats to use
- how to demodulate the signal
- how to achieve synchronization.

The main advantages of this type of transmission are:

- The power spectral density (power per unit of bandwith) is much lower because of the larger bandwith used. This implies lower interception probability and lower level of interference.

In order to recover the data, the receiver has to know the chip sequence, this increases the privacy level of the communication

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- The receiver restores the data to the original bit rate (which is much smaller than the chip rate). Therefore, the signal bandwith is reduced.

- Several communications can share the same spread spectrum, as long as they use uncorrelated chip sequences (as in the case of CDMA); mutual interference can be kept at a controlled low level.

In the 802.11 standard, the chip sequence is public and equal for all users. Therefore, the spectrum can be shared between several systems at a low level of mutual interference.

Codes in DS Systems

As codes are periodic, a suitable value of the periodicity has to be determined.

Short codes facilitate a good synchronization, while long codes are needed to guarantee the privacy of the communication.

Codes have to meet two important characteristics:

- A balance of "0"s and "1"s is needed, to avoid the presence of a DC component.

Usually pseudo-random codes are generated with equal probability of appearance of the two symbols.

- The chip rate has to be made a multiple of the bit rate. If so, both sequences are synchronized and the high-frequency energy resulting, if a transition in the bit sequence would be followed shortly after a transition in chip sequence, is minimized.

Codes can be categorized as:

- orthogonal
- quasi-orthogonal
- uncorrelated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. Orthogonal codes are codes for which the correlation in the same family of codes is null.

This is extremely useful because different users can use the same frequency band with theoretical no interference. These codes are obtained by using the same time reference. So the system has to maintain a perfect synchronization between al users. Orthogonal codes are used in the transmission from a central node to remote receivers.

An example of orthogonal code is the Walsh code used in the IS-95 standard.

2. Quasi-orthogonal codes can tolerate displacements at some degree, with no significant degradation of mutual interference.

They are usually used when a synchronization system is in use, by keeping the received signals aligned with a few chip intervals.

Subsets of the Gold codes are the quasi-orthogonal codes and are generally used in CDMA satellite communications.

Orthogonal and quasi-orthogonal codes are short codes, used to make synchronization easier.

3. Uncorrelated codes are kept in identical tables, both at the transmitter and at the receiver.

They are usually obtained from pseudo-random generation functions.

These types of codes are long codes, used to minimize the correlation between them.

Examples are maximal length codes and some type of Gold codes.

Transmission and reception in DS systems

Several modulation techniques are in use, like: FM, PSK, BPSK, QPSK.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

QPSK has the advantage of spreading the power with the same frequency between the two orthogonal carriers. This improves the performance of the system in the presence of narrowband interference signals.

BPSK has the advantage of an easier implementation and a higher immunity to noise.

A combination of the two modulation techniques is possible in order to combine the advantages, by using BPSK for data and QPSK for spreading sequence.

Performance in DS systems

The filtering process affects all type of signals, including noise, by reducing its power.

The performance of a DS receiver can be analyzed, by

- obtaining the total perturbation power within the spread spectrum bandwith, as:

$$P_{ss} = \Sigma I_i + N_{ss}$$

and then, by finding the signal to perturbation ratio in the narrowband bandwith, as:

$$\left(\frac{S}{P}\right)_{nb} = G_p \left(\frac{S}{P}\right)_{ss}$$

where:

$\left(\frac{S}{P}\right)_{nb}$ is the signal to perturbation ratio in a narrowband system

$\left(\frac{S}{P}\right)_{ss}$ is the signal to perturbation ratio in a spread spectrum system

G_p is the processing gain, defined as the ratio of chip rate to bit rate.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

According to the optimum perturbation theorem, the combination of all interfering signals and noise, after the correlation, is noise-like. Therefore, the relationship that relates the BER to the signal-to-noise ratio can be applied.

The main advantage of the DS transmission is that it allows a simultaneous extension of both the threshold and the processing gain.

FHSS protocols and functions

The FHSS physical layer consists of two protocol functions:

- Convergence function provides the interface with the MAC layer and adapts the frame format to of the MPDU to the suitable format of the FH transmission.
- Medium-dependent function, defines the characteristics of the transmission/reception of data through radio.

2. MAC sublayer

The most important general requirements for an efficient MAC layer are:

1. Throughput

Is one of the most critical consideration of the MAC sublayer. The aim is that the capacity of WLAN to be as close as possible to that of the wired counterpart.

The initial target was to operate in the range of 1-20Mbps. In the Ethernet (802.3) 10Mbps standard the real operational throughput is only 3-3.5Mbps.

In WLAN a larger real throughput can be obtained by using spread spectrum techniques, which allow multiple transmissions simultaneously.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

An other important aspect of throughput is the impact of unauthorized access. Therefore MAC layer should contain a successful security scheme in order to minimize the impact of unauthorized users.

2. Low Delay

Delay is an important characteristic especially in WLANs conceived for multimedia applications, some of them of real-time use, in which case delay should be minimized.

3. Transparency to different PHY layers

WLANs and the 802.11 standard works with different physical transmission layers, like:

- Direct Sequence spread spectrum
- Frequency Hopping spread spectrum
- Diffuse Infrared.

These PHY layers are different both in system design and propagation characteristics. A way to achieve this is to conceive a MAC-PHY interface.

4. Fairness of Access

For different reasons (like the fading characteristics of channels) situations could come up in which one or some users have an unfair access to the network. Mac layer should be able to resolve such situations, being possible to get fair access at 6-9 dB power difference.

5. Battery power consumption

As wireless devices are mainly portable, they are typically battery-powered. Therefore, devices have to be extremely energy-efficient. Therefore very limited power should be used for packet transmission. Sleep-mode should be used at user-end.

6. Maximum number of nodes and maximum coverage area

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A WLAN should support hundreds of nodes. Therefore, the MAC layer should not limit the maximum number of nodes to maintain a satisfactory performance.

The typical area coverage for a WLAN is in the range of 100m2, which introduces a propagation delay under 1milisec. This delay can cause major problem for some type of MACs like, synchronous CDMA.

7. Robustness vs. channel access and interference

Another big challenge is the ability of a designed WLAN to work successfully, in case of collocated networks. Generally there are two major concerns for collocated networks:

- Security: A wireless network is more difficult to secure since the transmission medium is open to anyone within the geographical range of the transmitter. Data privacy is usually insured through encryption. That is given at the cost of decreased performance and increased cost. An example of widely used encryption algorithm is Wired Equivalency Privacy (WEP), based on RC4 PZRNZG, developed by RSA Data Security.

- Interference: is caused by simultaneous transmission from two or more sources, sharing the same frequency band. Also, it could be caused by the so-called "hidden terminal" problem: a station cannot detect that a transmission already started. Interference is also cause by multipath fading, characterized by random amplitude and phase fluctuation at the receiver.

The reliability of the communication channel is measured by the average BER. Packet loss in the range of 10-2 is generally acceptable. Automatic repeat request and forward error correction are usually used to increase reliability.

8. Multicell coverage

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Multicell coverage is accomplished by the use of Access Points (AP), named also Base Stations, and which are repeaters. It's necessary that the coverage of one cell to overlap adjacent cell properly: the overlapping region has to be minimized in order to increase system capacity, but at the same time, seamless service should be guaranteed.

9. Self-interference

When two different APs try to transmit packet to a node in the joint (overlapping) region simultaneously, this may cause interference and the loss of the packet.

10. Self-collision

Reverse- a node found in a joint region may transmit a packet that is received simultaneously by more than one AP. This will cause collision and bandwith waste, while the packet is routed to the destination.

11. Up-down collision

While a node A is transmitting up-link, another node B is receiving down-link. Node B may be able to "hear node A's transmission.

A possible collision window will always exist (situation can be alleviated but not eliminated).

12. Ad-hoc connectivity

Ad-hoc connectivity is a peer-to-peer connectivity without prior knowledge of the network topology

13. Hand-off roaming between service areas

In indoor areas, because of fast fading this is not a straightforward problem. For time-bounded services, a real-time hand-off is also a challenging issue. Also power consumption is could be a problem in this case.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

14. Multicast/broadcast

15. Insensitivity to capture effects

Although the capture effect can increase through put it can also prohibit fair access. A solution is to enforce insensitivity at the receiver end. This will enhance physical transmission and avoid potential problems from capture.

16. Priority and non-symmetrical traffic

Traffic with different priorities should be supported. Most of the situation, up-link and down-link traffics are not of the same size: non-symmetrical traffic should also be supported.

17. Human safety

Research is underway to determine if RF transmission is health related. Transmission power should be kept as low as possible and IR transmission should not impair vision, in order to avoid safety complications.

The Structure of the MAC Protocol

A WLAN working in the 802.11 standard consists of one or more Basic Service Sets (BSS) that are interconnected through a Distribution System (DS). A BSS is defined as a group of stations (STA) found under the direct control of a single coordination function. The coordination function could be:

- Distributed Coordination Function (DCF) is based on a contention scheme, Carrier Sense Multiple Access with Collision Avoidance (CSMA-CA), and

- Point coordination function (PCF) is based on a contention free scheme, being used in a network with time-bounded services, where the Point Coordinator (PCF Station) has

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

priority control over the medium. When PCF is active, it allows only one station from each cell to have priority access to the medium, at any given time.

The geographical area covered by a WLAN is named as Basic service Area (BSA), which is analogous to a cell in the cellular communications network.

From the type of communication point of view, we can distinguish 2 large categories of WLANs:

1. Ad-hoc network, in which any 2 stations from a BSS can communicate directly one with the other, without any additional infrastructure.

The main characteristics of a ad-hoc network are:

- lack of Access Points (AP)
- no functionality to support mobility
- support for data transfer only between stations belonging to the same BSS

2. Infrastructure network is established to use Access Point. The AP supports range extension, by providing connectivity between multiple BSS, thus forming Extended Basic Service Set (ESS). An ESS is made up of multiple BSS using a common DS. Therefore DS can be understood as a backbone network, and it's implementation independent.

We note that although DS could be physically the same as BSS, we have to treat it logically as different, as being used as a transport medium between different BSS from the same ESS. DS also could be a gateway to wireless-wireline connection.

The gateway in this case is named a portal, having functionalities similar to a wireline bridge; it provides range extension and transfer between different frame formats. A DS is implementation independent; therefore it could be another wireless medium or as well a wireline medium.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The 802.11 standard of WLAN supports one mandatory and two optional coordination function schemes, as follows:

1. Distributed Coordination Function (DCF): is a mandatory coordination function scheme, based on a Carrier Sense Multiple Access (CSMA) with Collision Avoidance CA, and notated CSMA/CA.

2. A variant named DCF with handshaking, is an optional alternative, in which two additional frames are used:

- Request to send (RTS), and
- Clear to Send (CTS).

The use of these two additional frames limits the problem of hidden station.

3. Point Coordination Function (PCF) for Distributed Time-bounded Services (DTBS), in which a Point Coordinator (PCF) has priority control of the medium, i.e. when PCF is active, it allows to only a single station in a cell to have exclusive access over the medium., at any one time.

Further, we will describe in more detail these 3 types of schemes:

1. Distributed Coordination Function (DCF)

DCF is the fundamental access method, used to support asynchronous transfer on a best effort basis. It's a mandatory method, so all stations must support it.

DCF operates only in an ad-hoc network and works either solely or in cooperation with PCF in an infrastructure network.

DCF supports contention services, i.e. each station with an MSDU queued for transmission must compete for access to the channel and once the MSDU is transmitted, the race for the channel start all over again.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Contention services promote a fair access to the channel, for all stations.

DCF is based on CSMA/CA. CSMA is a member of the ALOHA family of protocols.

In fact represents the simplest way to improve ALOHA.

CSMA is applied both to wired and wireless networks, with the difference that in wired network, Collision Detection (CD) is used, instead of CA.

IN 802.11 Carrier Sensing is performed both at physical layer level (PHY) as well as at MAC level, in which case is named virtual carrier sensing.

Physical carrier sensing detects the presence of other users by analyzing all detected packets and detects activity in the channel from other sources, via relative signal strength.

CA is used instead of CD because in this case a station cannot listen to the channel for collisions, while transmitting. If CD would be used, in case of a hidden node, a receiving station can hear two different transmitters, but these two transmitters cannot hear the transmission of one another. This could generate collisions.

Usually CSMA/CA is associated with polling or handshaking, because multipath fading (especially for indoor channels) lasts for the duration of a symbol.

2. CDF with handshaking (RTS-CTS)

In this case, in order to solve the hidden node problem, the protocol makes use of two control frames:

- Request to Send (RTS) frame is sent by a transmitter to a receiver

- A Clear to Send (CTS) frame is sent in response to the transmitter's RTS frame, giving permission to the transmitter to transmit and notifying all other stations to not initiate any transmission for a period of time. The CTS frame is also named Net Allocation Vector

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(NAV). NAV indicates the amount of time before the current session transmission is complete, and the channel can be sampled again for idle status.

Because of the larger overhead, RTS/CTS is not used for short packets, because in this case the likelihood of collision as well as the cost of retransmission is small.

The average peak throughput in this case is around 65% as opposed to simple CDF, when is around 80%.

The transmission procedure is:

A source station does a carrier sensing by sending MPDU duration information in the header of RTS, CTS and data frames. MPDU contains information, payload, and a 32-bit CRC. The duration field indicates the amount of time (in microseconds) in which, after the transmission of this frame, the channel will be used to transmit data frame or management frame.

The priority access to the medium is controlled using Inter-frame Space (IFS), which is a time interval between two frames, when no transmission happens.

There are three types of IFS:

- Short IFS (SIFS): is the smallest in size, and gives the highest priority access to the medium

- Point Coordination Function IFS (PIFS): is of medium size, and gives medium level access to the medium

- DCF IFS (DIFS): is larger in size, and gives the lowest priority level access to the medium.

Stations having SIFS have priority access to the medium over those stations required to wait PIFS or DIFS.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

When a station senses that the channel idles, it will transmit a MPDU.

The receiving station checks for the integrity of the transmission, waits SIFS time interval, and then it sends an ACK frame to the source station, which indicates that the transmission was successful.

When the data frame is transmitted, its duration field will let know to all other stations in the BSS, for how long the medium will be busy. This time is a sum of the duration for data transmission, for ACK transmission and for SIFS. All stations hearing the data frame will adjust their own NAV, based on the duration of the field value.

- The maximum data frame size is 2,346 octets;
- The RTS frame is only 20 octets, and
- The CTS frame is 14 octets.

The source station, after successfully competing for the channel, first it transmits the RTS control frame, followed by the CTS frame, sent by the receiver. All stations from the BSS, will update their own NAV based on the RTS received from the transmitter and CTS received from the receiving station. After receiving the CTS signal, the transmitter is sure that the channel is stable and reserved for a successful transmission of the MPDU. After each MPDU is successfully received, the receiving station sends a ACK frame to the transmitting station.

In case a collision happen with either the RTS or CTS frame, far less bandwith is wasted, as compared with a collision with a data MPDU.

Stations set to CSMA/CA with RTS/CTS can work in one of the following ways:

- to not use RTS/CTS
- to use RTS/CTS only after the RTS_Threshold value is exceeded

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- to always use RTS/CTS.

In a lightly loaded medium, where collisions a scarce, the use of RTS/CTS frames is a waste of bandwith.

DCF collision avoidance procedure

Collision avoidance is carried out through a random back-off procedure. If the medium is busy, the transmission candidate station is looking for a DIFS period. When this occurs, the station backs-off for an additional, randomly generated period of time before competing for the channel. This randomly waiting period of time is named Contention Window (CW), used to minimize collision during contention between multiple stations. The back-off timer is decreased only when the medium is idle, for at least DIFS period of time.

The back-off interval is chosen using the formula

$CW = \text{Int}[2^{i+1} . \text{ranf}( )] . \text{Slot_Time},$

where:

- i is the number of consecutive times a station tries to send an MPDU
- ranf is a uniform random variable in the range (0,1)
- Slot_Time is a time period.

In standard 802.11, time is slotted in time periods that correspond to a Slot_Time, which is much smaller than an MPDU.

Slot_Time is different for each Physical channel implementation. It is determined experimentally, using the following formula:

Slot_Time = Rx_Tx . APT where,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Rx_Tx is the time to determine that the channel is (i.e. the time required to switch from receiving to transmitting mode and vice-versa).

- APT is the air propagation time.

The main advantage of this access mechanism is that it promotes fairness: each station have equal probability to win access to the channel after the DIFS interval of time.

The main disadvantage is that could not support DTBS: there is no mechanism to guarantee minimum delay to stations, so, it does not support time-bounded services, like voice transmission.

3. Point Coordination Function (PCF)

PCF is an optional capability; it is connection-oriented, and provides contention-free frame transfer. PCF relies on the Point Coordinator (PC) to perform polling for a transmission station and enabling the chosen station to transmit without competition for the channel. Usually the AP in each BSS is also a PC, but it could be also a specialized equipment. It is left to the implementers to determine the polling sequence and how is the polling table maintained.

The main applications using this type of transmission are time-bounded (like voice and/or video) in which packet order is a main concern. Generally, PCF have to coexist in the same BSS with DCF.

Stations capable to work during a Contention Free Period (CFP) are named CF-aware stations. A CFP Rate is used to determine the frequency with which PCF occurs.

The remaining part of the time is used for contention-based traffic.

The CFP interval starts with a beacon frame transmitted by the AP/PC. The beacon is also used for synchronization and timing. The maximum value of the CFP is a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

manageable parameter CFP_Max_Duration. It varies between the minimum time required to transmit two maximum size MPDU, including overhead and a maximum.

It is up to the AP/PC to decide how long to operate the CFP during any repetition interval: if the PCF type of traffic is light, it may shorten the CFP and allocate the remaining time to DCF type of traffic.

The maximum amount of delay for a frame on a CFP interval is the time needed to transmit the RTS/CTS handshake, a maximum length MPDU, and the ACK signal.

During a CFP time interval, a station is allowed to transmit only in response to a poll from the Point Coordinator, or for the transmission of an ACK.

During a Contention Free Period (CFP) te following sequence of steps happen:

- At the start of a CFP period, the Point Coordinator (PC) senses the medium. If the medium remains idle for a PIFS interval, PC transmits a beacon frame to initiate the CFP.

- Then PC waits for SIPS time to transmit a CF-Poll, Data or Data+CF-Poll frame.

- In case there is no more request for CFP type of transmission, PC can terminate this type of transmission by transmitting a CF-End frame.

- Otherwise, if a CF-aware station receives a CF-Poll frame from the PC, then the station can respond to the PC after a SIFS idle-period, with a CF-ACK)no data) or a Data+CF-ACK frame.

- If PC receives a CF-ACK frame from a station, then PC can send a Data+CF-ACK+CF-Poll frame to another station. The CF-ACK portion is used to acknowledge the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

receipt of the previous data frame. The ability to combine polling and data and acknowledgement frames, gives higher efficiency.

- If PC sends a CF-Poll frame and the station has no data to send, it will send a Null-Function frame back, instead.

- If PC does not receive a ACK for a transmitted data frame, then it waits PIFS amount of time, and continues transmitting to the next station of the polling list.

- If a station receives a CF-Poll frame, it may choose to transmit a frame to another station in the BSS.

- When the destination station receives the frame , a DCF-ACK is returned to the source station, and PC waits for PIFS amount of time following the ACK frame, before transmitting any other frame.

- PC may also choose to transmit a frame to a non-CF-aware station. In this case if the station had a successful reception of the frame, will wait SIFS time and reply to the PC with a standard contention-period ACK frame.

- Fragmentation and reassembly is accommodated through a Fragmentation_Threshold value, used to determine whether MSDU are fragmented prior to transmission. It is the responsibility of the receiving station to reassemble the fragments of the original MSDU.

8.5. IEEE 802.11 Sub-standards

Next the most relevant 802.11 sub-standards are presented:

8.5.1. EEE 802.11a

The 802.11a standard was designed to operate in the more recently allocated 5GHz UNII

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Unlicensed National Information Infrastructure) band. It's not using spread-spectrum technology, but coded orthogonal frequency division multiplexing (COFDM), considered more suitable for indoor environments.

COFDM has superior performance to that of spread-spectrum technologies by breaking one high-speed data carrier into several lower-speed subcarriers, which are then transmitted in parallel. Each high-speed carrier is 20MHz wide and is broken up into 52 subchannels, each approximately 300KHz wide. COFDM uses 48 of these subchannels for data, and the remaining 4 are used for error correction. COFDM delivers higher data rates and a high degree of multipath reflection recovery because of its encoding scheme and the type of error correction.

Each subchannel in the COFDM implementation is about 300KHz wide. BPSK (binary phase shift keying) is used to encode 125 Kbit/s of data per channel, resulting in a 6 Mbit/s data rate. By using quadrature phase shift keying (QPSK), the amount of data encoded can be doubled to 250 Kbit/s per channel, resulting an overall 12 Mbit/s data rate. Additionally, by using 16-level quadrature amplitude modulation encoding, 4 bits per hertz, a data rate of 24 Mbit/s can be achieved. The 802.11a standard specifies that all 802.11a-compliant products must support these basic data rates. The standard also allows vendors to extend the modulation scheme beyond this 24 Mbit/s. But, if a larger amount of bits are encoded per cycle (Hz), than the signal becomes more susceptible to interference and fading, resulting a shorter useful range of transmission. The range still could be increased by increasing the output power.

Compared to 802.11b, because transmission happens at a higher 5GHz frequency, the transmission distance will decrease, if the radiated power and encoding scheme are not

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

changed. The 802.11a standard overcomes some of the distance loss, by increasing the EIRP to the maximum 50 milliwatts, in the first 100MHz of its allocated frequency band. **802.11a** standard also provides additional channel support. If with 802.11b, three channels are available for simultaneous operation (in the 2.4 to 2.4835GHz range), 802.11a supports up to eight channels operating simultaneously in the two lower bands of the 5GHz spectrum (in the USA is 5.15-5.25GHz and 5.25-5.35GHz. The high band of the unlicensed 5GHz spectrum (5.725 to 5.825GHz), is also available, but is more commonly used for building-to-building wireless applications. Regardless of the number of available channels, an 802.11a access point only uses one channel at a time. This creates an opportunity for vendors to develop load-balancing technologies to automatically spread the load fairly evenly between channels, as no-one becomes overcrowded.

802.11a has an overall bandwith of 54Mb/s of which 22-26Mb/s are for proper data transmission and the remaining used for overhead signaling. In fact, The specification also allows for the implementation of higher speed modes, for increased bandwidth. Alternatively named "turbo" mode or "2X" mode, this is expected to increase available bandwidth by between 25% to 50%. The disadvantage of the turbo mode is that it's not strictly specified and could lead to incompatibilities between implementations.

802.11a can operate at different speed levels. Giving the possibility to auto-adjust speed level depending on the strength of the signal. As a result in 802.11a one can operate at a maximum rated speed of 54 Mbit/s but can drop back to 48, 36, 24, 18, 12, 9, and 6 Mbit/s.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The two essential advantages of 802.11a standard over the 802.11b are: additional bandwidth and freedom from interference.

Since 802.11a operates in the 5GHz frequency range, the problems associated with an overcrowded frequency band common to 802.11b wireless networks does not exist at the moment. This may change if the 5GHz bandwith will be used in the future for other appliances, as, for example, cordless phones.

The main disadvantage of 802.11a standard is that it is not directly compatible with the already widely used 802.11b standard. This can be solved either by using bridging, or by deploying network elements working in both standards.

Standards bodies in the USA and Europe were unable to agree, as of date, on the 802.11a standard (renamed to Wi-Fi5 by WECA), which could lead to a fragmentation of WLAN market (similar to what happened to the mobile telephony).

The result of this disagreement lead in Europe to the adoption of a variant of the 802.11a standard, named IEEE 802.11h.

8.5.2. IEEE 802.11h: addresses the concerns that 802.11a could potentially cause interference with military satellite transmissions. It solves it, by adding support for TPC (Transmit Power Control), by:

- preventing a PC card from emitting more radio signal than is needed, and
- introducing Dynamic Frequency Selection (DFS), which lets a device listen to what is happening in the airspace, before selecting a channel.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

8.5.3. IEEE 802.11b

The 802.11b group within IEEE was driven largely by Lucent Technologies and Intersil Corp.

It is designed to operate in the 2.4GHz ISM (Industrial, Scientific and Medical) band.

802.11b brought to important contributions to the WLAN standards that gave the possibility of increasing transmission speed to 5.5Mb/s and 11Mb/s:

- choice of Direct Sequence Spread Spectrum (DSSS) for physical layer transmission. (Frequency Hopping (FH) could not support the higher speeds without violating current FCC regulations)

- choice of a more efficient coding scheme, known as complimentary code keying (CCK), giving the possibility to increase transmission to a top-end data rate of 11Mbit/s.

802.11b systems will interoperate with 1Mbit/s and 2Mbit/s 802.11 DSSS systems, but will not work with 1Mbit/s and 2Mbit/s 802.11 FHSS systems.

Interoperability problems were effectively eliminated by setting up the Wireless Ethernet Compatibility Alliance (WECA). WECA's interoperability testing and certification program plays an important role in eliminating proprietary or dead-end solutions.

802.11b quickly became of a wide use first by the rapid increase in deployment of ad-hoc enterprise settings, in educational and institutional networks and more recently in home use, either by use of laptops or to share broadband Internet connections.

The most important disadvantages of the 802.11b standard consist in using the already crowded frequency band of 2.4GHz. Other devices using the same frequency band include microwave ovens, cordless phones, legacy wireless systems, home control

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

devices using the X-10 standard, and more recently by all devices using the Bluetooth standard.

802.11b operates in a range of 15-45 m indoors, and up to 300m outdoors.

Even if the nominal speed is 11Mb/s, overhead, configuration and security factors can reduce actual throughput to a typical 5Mbit/s. While that is sufficient for web browsing, it's inadequate for more demanding applications such as video streaming.

The Physical Layer Convergence Protocol (PLCP) preamble (used to determine transmission speed and assure synchronization) consists of 24 bytes, as opposed to only 8 bytes for wired Ethernet.

Obstacles in the path of signals can result in speed drop, in order to maintain connection Collision avoidance methods impose significant bandwidth overheads.

Security is another important issue: Wired Equivalent Privacy (WEP), part of the wider 802.11 standard, have vulnerabilities, at easy to decrypt the 40-bit security keys. In order to make hacking much more difficult, security keys that shift dynamically rather than remain static, were added, outside the standard. But these measures are not foolproof and make interoperability of different equipment, a problem. At this time there are conceived several standardized security solutions, like a new 128-bit encryption algorithm known as Advanced Encryption Standard (AES) that will require hardware or firmware upgrades, and a Temporal Key Integrity Protocol (TKIP) that will be compatible with WEP.

In addition to taking care to ensure that WEP is enabled, there are a number of other simple measures users can take, to increase their level of protection, like:

- changing all network names and passwords,
- using MAC (Media Access Control) address blocking, or

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- setting up the network as a closed system.

Security can be tightened even more by implementing a Virtual Private Network (VPN) or RADIUS authentication server.

802.11b does not have any Quality of Service (QoS) provisions for multimedia content.

802.11a and 802.11b standards were developed at the same time and published in 1999. It happened that 802.11b penetrated the market prior to 802.11a, which opened a relative advantage for 802.11b. But, on the other hand, some problems encountered in 802.11b, could be lately addressed and fixed in 802.11a.

8.5.4. IEEE 802.11g

In 2001, advantageous elements from both 802.11a and 802.11b were selected, in order to issue a compromise standard, coded 802.11g . Because it relies on the fundamental features from both standards, 802.11g is inter-operable with both 802.11a and 802.11b. This will make easier the migration from any of the two previous standards.

So, the main advantage of 802.11g is that it provides a clear compatibility between the 802.11a and 802.11b standards and encourages the future development of true multi-mode/RF devices.

The main disadvantage is that the already congested 2.4GHz frequency band will get even more crowded. On the other hand, it's also true that one of the major reasons why the 802.11b standard enjoyed international acceptance was because the 2.4GHz band is almost universally available, and where there are conflicts vendors can implement frequency-selection software that prevents a radio from operating at illegal frequencies. By contrast, the 5GHz spectrum does not have this advantage at the time being. Because

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

parts of the 5GHz band are used by military applications, such as high-energy radar, has resulted in placing regulatory restrictions on the commercial use of the band. For example in Japan only the lower 100MHz of the frequency spectrum is available, which means 802.11a applications in Japan will face more contention. In Europe, the lower 200MHz are common with the FCC's 5GHz allotment, but the higher 5.725GHz to 5.825GHz band reserved for outdoor applications are occupied. Even in the USA, where 802.11a is more widely used, there are questions about security risks for military operations.

As in in the Case of the wired Ethernet case, the faster standard is not expected to become quickly generalized, until bridging network elements become available.

8.5.5. IEEE 802.11e

The IEEE 802.11e specification is believed to become the industry's first true universal wireless standard, by offering seamless interoperability between business, home and public environments (such as airports and hotels), and at the same time, to meet the unique needs of each.

Unlike other wireless standards, 802.11e is the first one that reaches at the same time home and business environments. Also, it adds quality-of-service (QoS) features and multimedia support and it's backward compatible with the existing 802.11b and 802.11a wireless standards.

QoS and multimedia support are critical to future wireless networks where voice, video and audio will be delivered. With QoS new services, available at this time only in wired

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

networks, video on demand, audio on demand, voice over IP and high-speed Internet access, will become available for wireless users too. QoS is also a critical element for consumer electronic companies looking to offer home wireless networking devices.

The QoS Baseline proposes a better way to handle time-sensitive traffic for multimedia applications. The QoS Baseline accommodates both time-scheduled and polled communications during contention-free periods (CFP). It also offers improvements in the efficiency of polling. The QoS Baseline also proposes improvements in channel robustness, which is achieved through forward error correction (FEC) and selective retransmission. It also covers adaptable stream-service interfaces for admission control in higher layers and mechanisms for dealing with adjacent subnets operating on the same wireless channel.

Improved channel access during CFP and the ability to retain polling for backward compatibility result in more efficient polling. It also contains the ability to schedule transmissions, and chain a sequence of polls in a single command.

These mechanisms maximize the efficiency of high-bandwidth streams, provide power-management friendly implementations, and polled-style access for variable bit rate and bursty streams.

Channel-access methods can be used in any combination by point coordinators in order to enable devices to access the channel, during CFP.

A centralized scheduler is conceived in the QoS Baseline to guarantee collision avoidance and, consequently, to improve the ability to deliver time-critical payloads. The standard implements the ability to handle critical QoS constraints like latency, jitter and bandwidth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

System reliability is achieved by tying channel access to the allocations made by subnet bandwidth manager (like higher-layer protocols).

In addition, channel robustness becomes an important consideration in 802.11e, because noise, interference and multipath effects lead to degraded channel throughput in the 2.4-GHz and 5.x-GHz bands, adversely affecting the ability to reliably transmit latency-sensitive or high-bandwidth traffic, such as voice and video. The proposed schemes include: FEC and selective retransmission. These mechanisms include the ability to specify the correction, acknowledgement and retransmission policy on a per-stream basis, thereby accommodating a range of traffic types with policies designed specifically for each.

Channel throughput is improved through forward error correction (FEC), delayed acknowledgements with selective retransmissions, and dynamic channel change.

These enhancements will support QoS even when wireless subnets are deployed densely, like in an enterprise environment, where subnets could be located within radio range of each other and, otherwise, could cause interference and/or collisions during the communications by devices situated in different subnets.

802.11e is still in draft mode at the time being.

8.6. HiperLAN2

HiperLAN2 is one of the new generation standards supporting both asynchronous data communication and time critical communication services, like packetized voice and video.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It is being developed by the European standard body ETSI, under the general project name of BRAN (Broadband Radio Access Networks).

The HiperLAN2 standard is very close to 802.11 in terms of its physical layers: both use OFDM technology to achieve their data rates, but is very different at the MAC (Media Access Control) level and in the way the data packets are formed and devices are addressed.

More precisely, if 802.11 can be viewed close to Ethernet, HiperLAN2 is much closer to a wireless Asynchronous Transfer Mode (W-ATM).

HiperLAN2 operates by sharing the 20MHz channels, in the 5GHz spectrum, in time, using Time Division Multiple Access (TDMA) to provide QoS through ATM-like mechanisms. This is in fact considered as the main advantage over 802.11h, the standard it is compared to.

HiperLAN2 is conceived as a cellular networking topology combined with ad-hoc networking capability.

It supports two basic modes of operation: centralized mode and direct mode:

- The centralized mode is used in the cellular networking topology where each radio cell is controlled by an access point covering a certain geographical area. In this mode, a mobile terminal communicates with other mobile terminals or with the core network via an access point. It is mainly used in business applications - both indoors and outdoors - where an area much larger than a radio cell has to be covered.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- The direct mode is used in the ad-hoc networking topology - mainly in typical private home environments - where a radio cell covers the whole serving area. In this mode, mobile terminals in a single-cell home "network" can directly exchange data. HiperLAN2 and 802.11h networks do not interfere in their operation but they are not compatible.

Bibliography

8.1. Asuncion, Santamaria, F.J. Lopez-Hernandez. "Wireless LAN: Standards and Applications". Artech House. 2001.

8.2 Gast, M. "802.11 Wireless Networks: The Definitive Guide". O'Reilly and Associates. 2002.

8.3 Geier, J. "Wireless LANs", Second Edition. SAMS. 2001.

8.4. IEEE 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification". 1997.

8.5. IEEE 802.11a "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High Speed Physical Layer Extension in the 5GHz band". 1999.

8.6. IEEE 802.11b "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High Speed Physical Layer Extension in the 2.4GHz band". 1999.

8.7. IEEEE. "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications". IEEE Standard 802.11, IEEE. June 1999.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

8.8. http://www.wi-fiplanet.com/tutorials/article.php/2236611

8.9. http://www.hiperlan2.org.

8.10. Keller, R. "Wireless ATM for Broadband Multimedia Wireless Access: the ATM mobile project". IEEE Personal Communications, pg. 66-80. Vol.6 No.5, Oct. 1999.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 9. Mobile IP

Mobile IP is an open standard, defined by the Internet Engineering Task Force (IETF) in the RFC 2002, for wireless network Wide Area Network (WAN) environment. It allows seamless roaming between different IP networks by keeping same IP address. It is also highly scalable and compatible with any other IP type WAN, wireless or wireline.

Even if mobility could be implemented at data link-layer level, data crossing networks or different link layers, it still represents a problem. Mobil IP is addressing this issue, making different type of wireless networks capable to intercommunicate seamlessly.

A device on a network is reachable through normal IP routing by the IP address it is assigned on the network. But in case of a wireless network, where roaming is very probable, a problem occurs when a device roams away from its home network and is no longer reachable using normal IP routing. This would result in the active sessions of the device being terminated. Mobile IP was created to enable users to keep the same IP address in the case of moving to a different network without sessions or connections being dropped.

Because the mobility functions of Mobile IP are performed at the network layer, rather than the physical or data-link layer, the mobile device can travel through different types of wireless and wireline networks, while maintaining connections and implicitly accessed sites or applications. A variety of remote services, like remote login, remote printing, and file transfers, become possible by using Mobile IP. Also, certain network services, as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

software licenses and access privileges, are based on a static IP address and changing the IP address could compromise those services.

9.1. Components of a Mobile IP Network

Mobile IP has the following three components, as shown below.

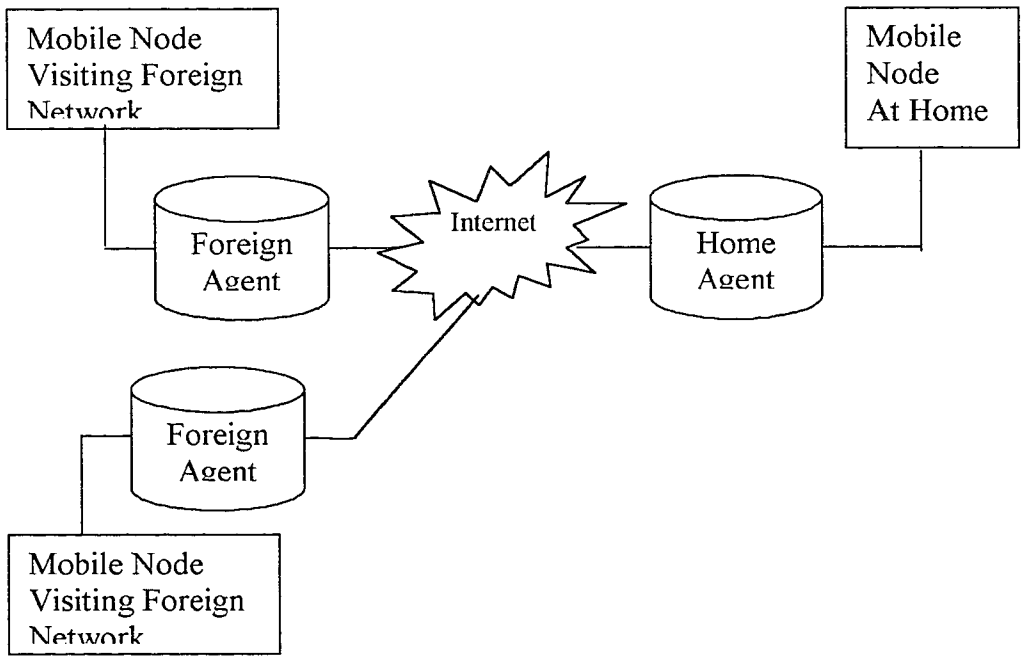


Figure 9.1 Mobile IP Components

1. **The Mobile Node** is usually a laptop or a PDA but it could be as well a cell phone with network roaming capabilities.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. **The Home Agent** is a router on the home network serving as the connection point for communication with the Mobile Node; it tunnels packets from a device on the Internet, called a Correspondent Node, to the roaming Mobile Node.

3. **The Foreign Agent** is a router that may function as the point of connection for the Mobile Node when it roams to a foreign network, delivering packets from the Home Agent to the Mobile Node.

The **care-of address** is the termination point of the tunnel toward the Mobile Node when it is on a foreign network. The Home Agent maintains an association between the home IP address of the Mobile Node and its care-of address, which is the current location of the Mobile Node on the visited network

9.2. Phases of Mobile IP

The Mobile IP process has three main phases:

1. Agent Discovery: a Mobile Node discovers its Foreign and Home Agents

2. Registration: the Mobile Node registers its current location with the Foreign Agent and Home Agent

3. Tunneling: a reciprocal tunnel is set up by the Home Agent to the care-of address (current location of the Mobile Node on the foreign network) to route packets to the Mobile Node as it roams.

A more detailed description of each phase follows:

9.2.1 Agent Discovery

During the agent discovery phase, the Home Agent and Foreign Agent advertise their services on the network by using the ICMP Router Discovery Protocol (IRDP). The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Mobile Node listens to these advertisements to determine if it is connected to its home network or to a foreign network.

The IRDP advertisements carry Mobile IP extensions that specify:

- whether an agent is a Home Agent, a Foreign Agent, or both;
- its care-of address;
- the types of services it will provide such as reverse tunneling and generic routing encapsulation (GRE); and
- the allowed registration lifetime or roaming period for visiting Mobile Nodes.

As an alternative, instead of waiting for agent advertisements, a Mobile Node can send out an agent request, named solicitation. This request forces any agents on the link to immediately send an agent reply, named advertisement.

If a Mobile Node determines that it is connected to a foreign network, it acquires a care-of address. There are two types of care-of addresses:

- Care-of address acquired from a Foreign Agent: is an IP address of a Foreign Agent that has an interface on the foreign network being visited by a Mobile Node. A Mobile Node that acquires this type of care-of address can share the address with other Mobile Nodes.

- Colocated care-of address: is an IP address temporarily assigned to the interface of the Mobile Node itself; it represents the current position of the Mobile Node on the foreign network and can be used by only one Mobile Node at a time.

When the Mobile Node hears a Foreign Agent advertisement and detects that it has moved outside of its home network, it begins registration.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

9.2.2 Registration

The registration is the most complex part of the three steps process.

The Mobile Node is configured with the IP address and mobility security association, which includes the shared key of its Home Agent. In addition, the Mobile Node is configured with either its home IP address, or another user identifier, named a Network Access Identifier (NAI).

The Mobile Node uses this information as well as the information it gets from the Foreign Agent advertisements, and form a Mobile IP registration request.

Registration implies the following steps:

1. The Mobile Node sends the registration request to its Home Agent. This may happen in two ways:

- either through the Foreign Agent, or

- directly, in case it's using a co-located care-of address and is not required to register through the Foreign Agent.

If the registration request is sent through the Foreign Agent, the Foreign Agent checks:

- the validity of the registration request, including lifetime limitations,

- the requested tunnel encapsulation is available, and

- that reverse tunnel is supported.

If the registration request is valid, the Foreign Agent adds the visiting Mobile Node to its pending list and then relays the request to the Home Agent.

If the registration request is not valid, the Foreign Agent sends back to the Mobile Node a registration reply with appropriate error code.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. The Home Agent checks the validity of the registration request, which includes authentication of the Mobile Node. If the registration request is valid, the Home Agent creates:

- a so called mobility binding, which is an association of the Mobile Node with its care-of address,

- a tunnel to the care-of address, and

- a routing entry for forwarding packets to the home address through the tunnel.

3. The Home Agent then sends a registration reply to the Mobile Node through the Foreign Agent, if the registration request was received via the Foreign Agent, or otherwise directly to the Mobile Node.

If the registration request is not valid, the Home Agent rejects the request by sending a registration reply with an appropriate error code.

4. The Foreign Agent checks the validity of the registration reply, including the fact that the associated registration request exists in its pending list.

If the registration reply is valid, the Foreign Agent:

- adds the Mobile Node to its visitor list,

- establishes a tunnel to the Home Agent,

- creates a routing entry for forwarding packets to the home address, and

- it relays the registration reply to the Mobile Node.

5. As a final step, the Mobile Node checks the validity of the registration reply, by checking that an associated request is in its pending list and that there is a proper authentication of the Home Agent.

If the registration reply is not valid, the Mobile Node discards the reply.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

If the registration is valid and it is accepted, then the Mobile Node is confirmed and all Mobility Agents are aware of its roaming. In the case of a colocated care-of address, it adds a tunnel to the Home Agent. Subsequently, it sends all packets to the Foreign Agent. The Mobile Node has to re-register before its registration lifetime expires. The Home Agent and Foreign Agent update their mobility binding and visitor entry, respectively, during re-registration.

If the registration is denied, the Mobile Node makes the necessary adjustments, depending on the cause of denial, and attempts to register again. For example, if the registration is denied because of time mismatch and the Home Agent sends back its time stamp for synchronization, then the Mobile Node adjusts its own time stamp for future registration requests.

So, in case of a successful registration a routing mechanism is provided in order to transport packets to and from the Mobile Node, as it roams.

9.2.3 Tunneling

The Mobile Node sends packets using its home IP address, giving the sensation to other Mobile Nodes that it is always on its home network, even when it's roaming on foreign networks.

Data packets addressed to the Mobile Node are initially routed to its home network, where the Home Agent now intercepts and tunnels them to the care-of address of the Mobile Node. Tunneling has two primary functions:

- encapsulation of the data packet to reach the tunnel endpoint, and
- decapsulation when the packet is delivered at that endpoint.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The default tunnel mode is IP Encapsulation within IP Encapsulation. Two other options are: GRE and minimal encapsulation within IP.

Typically, the Mobile Node sends packets to the Foreign Agent, which routes them to the communicating Mobile Node.

However, this data path does not solve fully the mobility issue because it does not reflect the true source for the data; it only reflects the home network of the Mobile Node.

Because the packets show the home network as their source inside the foreign network, and because of ingress filtering, the packets are dropped instead of being forwarded.

A feature, called reverse tunneling, solves this problem. In reverse tunneling, the Foreign Agent tunnels the packets received from the Mobile Node, back to the Home Agent.

Reverse tunneling is graphically represented in Figure 9.2.

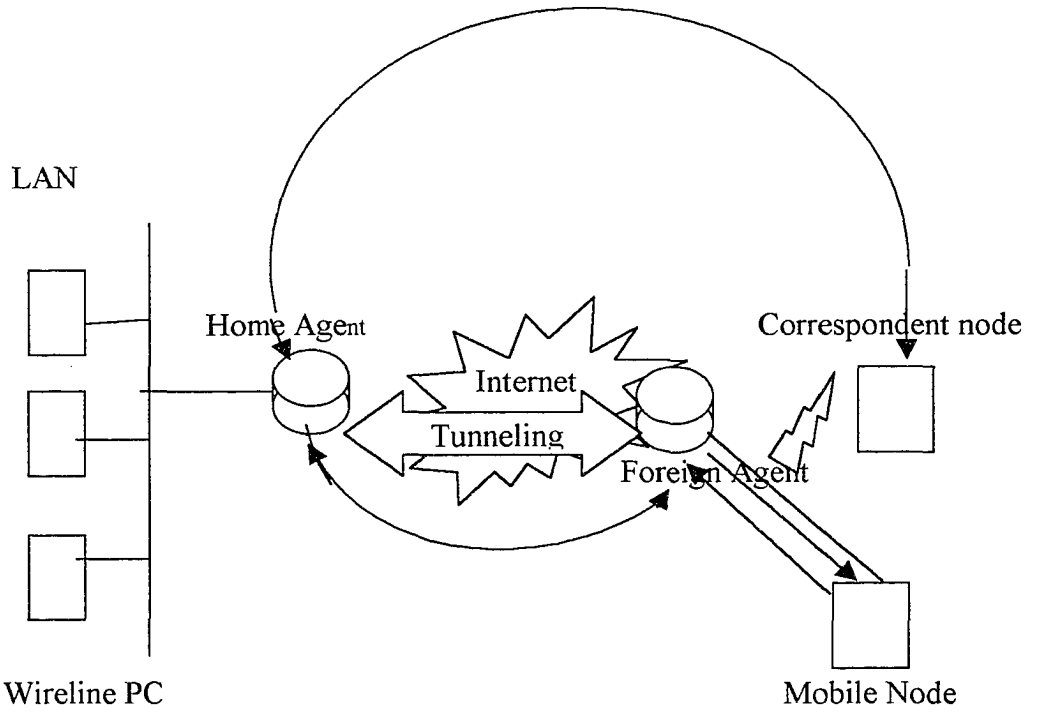


Figure 9.2. Reverse tunneling in Mobile IP

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

9.3 Mobile-IP Security issues

Mobile IP uses a strong authentication scheme for security purposes. All registration messages between a Mobile Node and Home Agent contain the so-called Mobile-Home Authentication Extension (MHAE).

For registration between a Mobile Node and Home Agent a 128-bit shared key is used.

To compute the authenticator value in the mandatory appended MHAE, the so-called keyed message digest algorithm 5 (MD5), in "prefix+suffix" mode, is used.

Mobile IP also supports the hash-based message authentication code (HMAC-MD5). The receiver compares the authenticator value it computes over the message with the value in the extension, in order to verify the authenticity.

Optionally, the Mobile-Foreign Authentication Extension (MFAE) and Foreign-Home Authentication Extension (FHAE) are appended in order to add stronger protection to message exchanges between a Mobile Node and Foreign Agent on one side and between a Foreign Agent and Home Agent, on the other side.

In the reply messages, the Home Agent returns its time stamp in order to synchronize the Mobile Node for registration.

Solutions to Mobility Offered by Mobile IP

It is estimated that Mobile IP provides a scalable, transparent, and secure solution to network mobility:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Mobile IP is scalable because only the participating components need to be Mobile IP aware, more precisely the Mobile Node and the endpoints of the tunnel. All other routers in the network, or the hosts with which the Mobile Node is communicating, are not aware of the movement of the Mobile Node between networks.
- Mobile IP provides mobility, transparently to any applications. In addition, the network layer provides link-layer independence, interlink layer roaming, and link-layer transparency.
- Mobile IP is secure because every time the set up of packet redirection is authenticated.

Bibliography:

9.1. "Encapsulation within IP" - RFC-2003. IETF. Oct. 1996.

9.2. http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/moblip/mobil_ip.pdf

9.3. Ioannidis, John, G.Q. Maguire Jr. "The Design and Implementation of a Mobile Internetworking Architecture". Proceedings of the Winter USENIX Conference". Jan. 1993.

9.4. "IPv4 Mobility Support" - RFC-2002. IETF. Oct. 1996.

9.5. Let, Hui, C. Perkins. "Ad-hoc networking with Mobile IP". Proceedings of the Second European Personal Mobile Communications Conference". Oct. 1997.

9.6. Perkins, Charles E. "Mobile IP, Design, Principles and Practices". Addison-Wesley. 1998.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 10. Multiservice Wireless Networks

From the invention of the telegraph and of the telephone we witnessed the emerging and maturing of three independent networks:

- Tele-communication network
- Data-communication network
- Voice/Video broadcast network

Each of these networks evolved and matured based on different principles, strictly related to the type of the carrying information. Each of them was fine-tuned depending on the type of the carrying information

With the advent of the computing industry and the deployment, not only of large data communication networks, but the interconnection of them into one huge worldwide network, the Internet, it became technologically feasible and economically sound to merge all these type of networks into one efficient type of network: Multiservice Network.

Multiservice Networks, (named also **Converged Networks)** represent the integration of several networks: data, voice, and video, into a single infrastructure, which initially will be a mixture of cell-based and packet based technologies and finally a truly packet based technology. Multiservice networking delivers the infrastructure needed to deploy new Web-based applications that represent a fundamental shift in how a business stays competitive and improves how a company interacts with customers, suppliers, employees, and industry partners. These applications include unified messaging,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

collaborative data sharing, video streaming and video conferencing, push-to-talk web pages, network-enabled call centers and other.

Infrastructure convergence does not mean only to replace the existing telephone networks. Today's telephone is very reliable but the transport of voice traffic could be much more efficient and newer technologies became available for integrated multimedia and web-based voice and video applications. The main reason for the generalization of Multiservice Networking resides in the capabilities of the new network to be ready for the web-based multimedia applications.

New required applications are the driving force for network advances. The new network applications have different operational requirements that demand different network services. Increased network traffic requires increased network bandwidth capacity, but new applications like IP telephony have other requirements. Increasing the network bandwith is not the only answer.

We can distinguish communication depending on rate type or delivery type:

a. Rate Type communications are:

a.1. Stream: is a predictable delivery at a relatively constant bit rate (CBR). For example,

although their rates often fluctuate, audio and video data streams are considered CBR because they have a quantifiable upper bound.

a.2. Burst: is an unpredictable delivery of "blocks" of data at a variable bit rate (VBR).

Applications, like file transfer, move data in bulk that can increase data rate to use all available bandwidth (no upper bound).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

b. Delivery Type communications are:

b.1. Asynchronous: has no constraints on delivery time (a.k.a. "elastic")

b.2. Synchronous: when data is time-sensitive, but flexible.

b.3. Interactive: when delays may be noticeable to users/applications, but do not adversely affect usability or functionality.

4. Isochronous: when it's time-sensitive to an extent that adversely affects usability.

5. **Mission-Critical:** when data delivery delays disable functionality.

The Demand for Convergence

Further, we present the main steps necessary for the convergence of the 3 well-known communication networks, toward a packet type network:

1. For data networks, the convergence to IP is basically already accomplished. Many information technologies and commerce applications have already moved to IP, and others are on the way, even if for various reasons some will always remain on legacy networks. Most applications are low-priority and low-bandwidth, with a high tolerance for delay, but others have strict operational requirements.

2. For radio and television networks, the convergence to IP has started but still has a long way to be accomplished. First of all, there is a need for much wider bandwidth. The broadcast models in existence today are a close match to the one-to-many IP multicast model. Broadcasters currently service millions of customers simultaneously, and unicast on the Internet could never possibly scale to these levels. Therefore, multicast deployment is necessary to enable the convergence of television and radio networks to IP.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. The added value that IP networks can provide for multimedia applications, is tremendous. They enable new dimensions to multimedia content. They can embed web links, or simultaneously deliver text, speech, sound, fix and mobile images. Also, they enable two-way communications, so content receivers can "talk-back" to providers. And since multicast allows receivers to talk to other receivers as well, they open the door to all-new application possibilities, like discussion groups, interactive distance-learning, game-shows, instant polling etc. They provide the potential for global audiences, which can make it economically feasible. But before this tremendous potential can be fulfilled, there are other network service requirements, like quality of service, that have to be implemented in addition to the multicast mechanism.

Real-Time and Two-Way Communication

IP telephony became today the most sought after application. However, its implementation revealed several deficiencies of IP type networks, and raised the need to define new standards and deploy managed bandwidth on IP networks. Even if voice communication is considered as a multimedia application, its bandwidth requirements are relatively modest (about 8Kbps each way), so it's not bandwidth the issue, but latency is a major one.

For IP telephony and other real-time or two-way applications, the timing requirements are much more significant than the bandwidth requirements. Dropouts and delays are noticeable and distracting and round-trip delivery delays above 500 milliseconds can make them unusable.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Because the TCP/IP protocol suite initially conceived for best-effort services, routing delays and lost packets due to transient network congestion, from the unpredictable and bursty nature of the network traffic, results in sub-optimal round-trip times on IP networks, and induce a serious limitation for IP-based telephone service.

The telecommunications industry started with a specific application, telephony, and built a network to suit it. Nevertheless, the Internet started exactly in the opposite way: it started with a new network technology and explored, successfully, new applications that were able to use the undefined, best-effort service.

Increasing bandwidth capacity will improve IP service, so it is an important first step towards solving the latency issue, but it is not enough to satisfy telephony application requirements.

Quality of Service (QoS) in Multiservice Networks

Quality of Service (QoS) can be defined as the ability of a network element (e.g. an application, host or router) to have some level of assurance that its traffic and service requirements can be satisfied. To enable QoS requires the cooperation of all network layers from top-to-bottom, as well as of every network element from end-to-end. Any QoS assurances are only as good as the weakest link in the chain between sender and receiver.

QoS does not create bandwidth but it manages bandwidth according to application demands and network management settings. Always, it cannot provide certainty if it involves sharing. Therefore QoS, with a guaranteed service level, requires resource allocation to individual data streams. The bandwidth allocated to an application in a "resource reservation" is no longer available for use by "best-effort" applications. But at

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the same QoS-enabled high-priority applications must not disable the low-priority applications. The worst case should be that low-priority applications simply have a lesser or slower service, but still have to function.

There are essentially two types of QoS available:

1. **Resource reservation** (integrated services): network resources are "reserved" according to an application's QoS request, and subject to bandwidth management policy. RSVP reservation protocol is one example of mechanisms to do this

2. Prioritization (differentiated services): network traffic is classified by priority and network resources are allocated according to bandwidth management policy criteria. To enable QoS, classifications give preferential treatment to applications identified as having more demanding requirements. The "Type of service" field from the IP protocol provides a rudimentary mechanism to do this. New, more sophisticated solutions were proposed or are under investigation.

These two types of QoS protocols and algorithms are not competitive or mutually exclusive, but on the contrary, they are complementary. As a result, they are designed for use in combination, in order to accommodate the varying operational requirements in different network contexts.

QoS implementation

There are several solutions to implement in an efficient way QoS:

a.) Protocols:

1. Implement new protocols for which QoS is inherent: like ATM protocol, or protocols able to define integrated services like RSVP reservation protocol or protocols able to define differentiated services like RTP/RTCP Real-Time

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Protocol/Real-Time Control Protocol; or specific protocols to voice/video integration like H.323, H.248 or SIP protocols

2. IPv4 has a field named "Type of Service" which allows a rudimentary implementation of QoS, named Type of Service
3. IPv6 gives the possibility to define "any-to-any" communication, a particular way of multicast communication, in a efficient way

b.) Hardware:

4. New network elements, like traffic shaper, give a better control of the delay variation (jitter) on the network
5. Multiservice switches and routers become more and more intelligent, provide switching at the upper layers of the OSI model, and give the possibility to better control traffic prioritization

c.) Software:

6. New filtering or bandwidth management software can "equalize" traffic and implement differentiated or integrated services, control and measure the QoS.

d.) Algorithms:

7. New and more efficient algorithms will be conceived and implemented at all levels of communications signaling, transmission, compression, authentication, traffic control, network management etc.

Virtual Private Network (VPN)

One of the main security tools in Multiservice Networks is the use of Virtual Private Networks (VPN). The essence of a Virtual Private Network (VPN) is its use of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Internet as a public data network. VPN extends the reach and capabilities of a company's network, in a very cost-effective way. A company can select the performance level, by choosing from a range of service options. Multiprotocol tunneling for transport client/server traffic provides a secure routing environment.

VPNs have the potential to solve many of the networking problems. VPNs allow network managers to connect remote branch offices and project teams to the main corporate network in an economic way and also to provide remote access to employees, while reducing the in-house requirements for equipment and support. Rather than depend on dedicated leased lines or frame relay's permanent virtual circuits (PVCs), VPN uses the open, distributed infrastructure of the Internet to transmit data between corporate sites. Companies using an Internet VPN set up connections to the local connection points, called points-of-presence or POPs, of their Internet service provider (ISP) and let the ISP ensure that the data is transmitted to the appropriate destinations via the Internet, leaving the rest of the connectivity details to the ISP's network and the Internet infrastructure. Because the Internet is a public network with open transmission of most data, Internet-based VPNs include measures for encrypting data passed between VPN sites, which protect the data against eavesdropping and tampering by unauthorized parties.

In addition, VPNs are not limited to fixed wireline users. As an added advantage, a VPN can provide secure connectivity for mobile users. These users can connect to their company's VPN by dialing into the POP of a local ISP, which reduces the need for long-distance charges and outlays for installing and maintaining large banks of modems at corporate sites.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Multiservice Network Management

Due to the convergence of voice, data and video, network management has become the connecting factor in the network. A wide range of network management solutions can be tailored to the customer demands, and which give the tools to operate and maintain a total quality network. At every stage in the network life cycle the network management system is the core, helping you to channel and maximize network control.

Multiservice Network Implementation

Migrating from yesterday's multiple networks into multiservice networks usually will happen in phases. Many companies only need to make minor changes to existing data network installations to enable voice integration. This may include adding multiservice-capable switches and routers to or may simply mean turning on the video and voice capabilities of an existing multiservice-capable router.

There is already a large consensus that the multiservice network integration will take place in two stages, probably first the infrastructure, and then the applications.

1. The infrastructure stage is often referred to as Voice IP Transport, meaning the existing Private Branch Exchanges (PBXs) providing voice service to end users remain in place and the converged network will now transport voice and data traffic between the PBXs or the Public Switched Telephone Network (PSTN).

2. The second stage, also known as IP Telephony (also named Desktop Telephony), involves IP Phones, voice-capable computer applications and Web-based multimedia applications that integrate voice and data at the desktop or local level.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Even if both stages are deployed simultaneously, the infrastructure piece is a key part of the planning process to ensure successful deployment of the applications.

Bibliography

10.1. Ahamed, Syed V., Victor B. Lawrence. "Intelligent Broadband Multimedia Network". Kluwer Academic Publishers. Boston, USA. 1997

10.2. Ahamed, Syed V., Victor B. Lawrence. "Design and Engineering of Intelligent Communications Systems". Kluwer Academic Publishers. Boston USA. 1997.

10.3 Ganz, Aura, Zvi Ganz, Kitti Wongthavarawat. "Multimedia Wireless Networks: Technologies, Standards and QoS" Prentice Hall. 2003

10.4 Jha Sanjay, Mahbub Hassan. "Engineering Internet QoS". Artech House. 2002.

10.5. "Voice and Fax over IP". At http://www.iec.org/online/tutorials/vfoip/index.html

10.6. Wesel, Ellen K. "Wireless Multimedia Communications". Adison-Wesley wireless Communications Series. 1998.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 11. Wireless VoIP

11.1. Voice over IP-based Networks (VoIP)

Internet Telephony also named Voice over IP (VoIP) or simpler, Voice over Net (VoN) deals with voice or fax communication services that are transported over the Internet, instead PSTN.

The basic steps imply:

- conversion of analog voice signal to digital format
- compression of the digitized signal
- translation into IP packets

at one end and the reversed process, the other end.

The first VoIP transmission was made in February 1995 when Vocaltec Inc. introduced its Internet Phone software. Conceived at that time to run on a 486/33-MHz PC and equipped with a sound card, speakers, microphone, and modem, the software compresses the voice signal and translates it into IP packets for transmission over the Internet. This PC-to-PC Internet telephony works, however, only if both parties are using the same Internet Phone software.

Since then, Internet telephony has advanced rapidly and soon will become the mainstream trend.

The main steps in VoIP development are:

- more and more companies are providing PC telephony software
- new interface equipment with PSTN become available; equipped with voice-processing cards, these gateway servers enable users to communicate via standard telephones and speeds up the integration of the two types of networks

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- new hardware and software technologies are developed in order to provide the quality of service necessary to transport the voice packets over the "best effort" IP network.

The process of creating voice IP packets is as follows:

- 1. The analog voice signal is converted to a pulse code modulation (PCM) digital stream, of 16 bits every 125 microseconds.

- 2. The line echo is removed from the digital stream. The resulted stream is further analyzed for silence suppression and tone detection.

- 3. The resulting PCM samples are transformed into voice frames and a vocoder compresses the frames. For example G.729 vocoder creates a 10milisecond voice frame containing 10 bytes of speech. Then it compresses 128 Kbps linear PCM stream into 8Kbps stream.

- 4. The voice frames are integrated into voice packets: first RTP packet with 12 byte header, then an UDP packet with 8 byte header and then an IP packet with 20 byte header is added.

- 5. The packet is sent through Internet, being routed by switches and routers, and finally delivered to the right destination.

- 6. At the destination the packet goes through a reverse process. Packets are assembled in the right order (even if not received in that order) to recreate the voice stream.

Internet telephony (VoIP) will have to accommodate already existing intelligent features from the telcom network like: call-waiting, three way calling, alternate billing, free call, emergency call, caller ID, voice mail and voice to text conversion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

11.2. QoS Issues for VoIP

The transport of the voice or video over packet networks raise some specific issues as delay, delay variation, and packet loss, presented below:

11.2.1 Delay

In the delivery of the voice, delay can be caused by various reasons, the most widely known being: voice digitization, voice processing, and voice transport.

Delay at its turn is causing mainly two problems: voice echo and talker overlap.

- Echo is caused by the signal reflections of the speaker's voice from the far-end telephone equipment back into the speaker's ear. Echo becomes a significant problem when the round-trip delay becomes greater than 50 milliseconds. As echo is perceived as a significant quality problem, voice-over-packet systems must address the need for echo control and implement some means of echo cancellation.

- Talker overlap is caused by a significant delay in the delivery of the voice signal to the recipient. It becomes significant if the one-way delay is greater than 250 milliseconds. The end-to-end delay is therefore the major constraint for reducing delay through a packet network.

The following are sources of delay in an end-to-end, voice-over-packet call:

- **Accumulation Delay (also named Algorithmic Delay)** is caused by the need to collect a frame of voice samples to be processed by the voice coder. Its magnitude is function of the type of voice coder used, and varies from a single sample time of 0.125 milliseconds, to many milliseconds.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- **Processing Delay** is caused by the actual process of encoding and collecting the encoded samples into a packet, for transmission over the packet network.

Combining the first two types of delays, we can define **encoding delay** as a sum of both the processor execution time and the type of algorithm used.

In some circumstances, it could be advantageous to group multiple voice-coder frames in a single packet, in order to reduce the packet network overhead.

- **Network Delay** is caused by the physical medium and protocols used to transmit the voice data and by the buffers used to remove packet jitter on the receiver side. Network delay is a function of the capacity of the links in the network and the processing that occurs as the packets transit the network. The jitter buffers, used to remove the packet-delay variation to which each packet is subjected, as it transits the packet network, in turn, could add additional delay.

11.2.2. Jitter

Jitter is a variation of the arrival time of packets at the receiver side after traversing the network. Jitter can be removed by collecting the arrived packets and holding them long enough to allow the slowest packets to arrive in time to be played in the correct sequence. But this will cause additional delay. This additional delay can be a significant part of the overall delay, as packet-delay variations can be as high as 70 to 100 milliseconds in some type of packet switched networks.

There are different schemes to solve the two conflicting goals of minimizing delay and removing jitter. The main goal is to minimize the size of the jitter buffer to a value which will prevent buffer underflow caused by jitter, but at the same time will minimize the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

additional delay of buffering. In other words the jitter buffer has to be sized in a way in which it reduces jitter without adding significant delay.

Two approaches to adapting the jitter buffer size are presented below. The choice is dependent on the type of network the packets are traversing.

- The first approach is to measure the variation of packet level in the jitter buffer over a period of time and incrementally adapt the buffer size to match the calculated jitter. This approach works best with networks that provide a consistent jitter performance over time, such as ATM networks.

- The second approach is to count the number of packets that arrive late and create a ratio of these packets to the number of packets that are successfully processed. This ratio is then used to adjust the jitter buffer to target a predetermined, allowable late-packet ratio. This approach works best with the networks with highly variable packet inter-arrival intervals, like IP networks.

11.2.3. Lost-Packet Compensation

Packet loss can happen in an internetwork transport for various reasons, mainly because of peak loads and congestions, temporary malfunctioning of some parts of a network, intentional or non-intentional shut down etc. In IP based network, which was initially conceived as a best effort service network, the incident of packet loss could be significantly higher than for example in ATM networks. The loss of voice packets cannot be treated the same way as data packets.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Below we present some schemes to minimize the impact of the packet loss in case of VoIP applications:

- Interpolation is a replacement of the lost packet with the last packet received during the interval when the lost packet was supposed to be played out. This is a simple method that fills the time between noncontiguous speech frames. It works well when the incidence of lost frames is infrequent, but it does not work well if there are a number of lost packets in a row or a burst of lost packets

- Redundant sending is a costly approach in which each voice packet is sent twice and on the receiver side each double is discarded. The method makes a precise correction of the lost packet but uses more bandwidth and also creates greater delay

- Hybrid approach is use of the Redundant sending method with low bandwith CODEC. This method will reduce part of the bandwith issue at the expense of the quality of the transmitted voice, and without solving the extra delay.

11.2.4 Echo Compensation

Echo is mainly caused by signal reflections generated by hybrid circuits that converts between a four-wire circuit (a separate transmit and receive pair) and a two-wire circuit (a single transmit and receive pair). The reflections of the speaker's voice are perceived by the speaker itself, therefore disturbing talk. Echo could be present even if no conversion of any kind happens. Always, this is acceptable if the round-trip delays through the network are smaller than 50 milliseconds. Also, this echo is generally masked by the normal side tone, every telephone generates.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Echo becomes a problem in voice-over-packet networks because the round-trip delay through the network is generally greater than 50 milliseconds. Thus, echo-cancellation techniques are always necessary.

ITU standard G.165 defines performance requirements that are currently required for echo cancellers. Also, G.IEC of the same ITU, defines much more strict performance requirements.

Echo is generated from the telephone network (circuit switched) toward the IP network (packet switched). The echo canceller is a digital filter which removes the echo from the telephone network, on the transmit path into the packet network.

11.3. Quality of Service solutions for VoIP

The most important feature of the converged networks should be the support of Quality of service (QoS) for the transfer of real-time information like: video or voice.

If the initial trend in voice data convergence was the use of ATM networks, several inconveniences like: high overhead of ATM cells, high cost, and the necessity to manage two different type of networks: IP and ATM, lead to a different approach. The actual trend is to merge all types of networks into IP type networks, the convergence being made at Layer 3.

Several standardization approaches were made by both the telcom and the datacom communities.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The main standards to manage data voice and video convergence are: H.323 and H.248 from ITU and SIP from IETF.

11.3.1. H.323 Specification

H.323.is an collection of protocols used to support audio, video and data applications.

It was conceived to support:

- convergent services across different types of networks

-point to point and multipoint conferences

- handle call control, multimedia applications and prioritization.

H.323 is an intelligent protocol but has scalability problems and does not interface directly with PSTN.

11.3.2. H.248 Specification

H.248 was initially designed at ITU and in a second stage was embraced with minor modifications by IETF under the name of MEGACO. H.248 uses a gateway/call agent architecture and support voice and video convergence with data transmission.

11.3.3 Session Initiation Protocol (SIP) was conceive by IETF as an application layer protocol, used to create, modify, and terminate session between two or more participants. The main applications SIP is intended to, are: multimedia conferences, internet telephony calls and multimedia distribution.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

SIP supports user mobility by proxying and redirecting requests to the user's new destination. SIP can be used to initiate sessions and to invite users to sessions, through other multicast protocols. SIP supports five type of multimedia communication initiation/termination:

- User location: determination of the end system to be used for communications

- User capabilities: determination of the media and media capabilities used for communication

- User availability: determination of the willingness to participate in a communication

- Call setup: establishment of call parameters at both ends and alerting the caller

- Call handling: transfer or termination of a call.

Designed by IETF is inherent to the Internet protocol suite and embeds more naturally into it. It cooperates, without being dependent on, with the following multimedia related protocols:

- RSVP (Reservation Protocol) used for reservation of network resources

- RTP (Real Time Protocol) used to transport real-time data

- RTSP (Real Time Signaling Protocol) used to control delivery of streaming media

- SAP (Service Advertising Protocol) used to advertise multimedia sessions via multicast

- SDP (Session Description Protocol) used to describe multimedia sessions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It is estimated that SIP will be the successful candidate in the generalization process of network convergence.

11.4. Voice Codecs

CODEC*s* perform the conversion from an analogue voice waveform to a digital stream of information.

The main steps are:

- sampling: a codec samples the value of the analogue signal at regular intervals (usualy every 125 microseconds, in order to comply with Niquist theorem, which states that a signal must be sampled at least twice as often as its highest frequency, in order to enable correct reconstruction of the signal when transmitted)
- quantising: the codec converts the measured analogue value into a numeric representation The resultant output is made up of discreet blocks of information, sent at regular intervals.

The bandwith is calculated assuming that voice samples can be transmitted in a datagram every 20ms. For coding algorithms which use much smaller sampling periods, multiple samples can be sent within each packet, and the samples can be buffered for up to 20ms. However, some algorithms do not produce samples which can be fitted exactly into 20ms (as G.723, which may reduce overhead).

The following tables shows the relevant characteristics of the most common coding algorithms:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Codec Name	Coding Algorithm	Sample (ms)	Bandwith (kbps)	IP bandwith (kbps)
G.711	PCM	0.125	64.0	80.0
G.723	ACELP	30.0	5.6/6.4	16.27/17.07
G.726	ADPCM	0.125	32.0	48.0
G.728	LD-CELP	0.625	16.0	32.0
G.729(a)	CS-ACELP	10.0	8.0	24.0

Table 11.1 Coding algorithms - main characteristics

where the coding algorithm acronyms are :

PCM - Pulse Code Modulation

ACELP - Algebraic Code-Exited Linear Prediction

ADPCM - Adaptive Differential Pulse Code Modulation

LD-CELP - Low-Delay Code Excited Linear Prediction

CS-ACELP - Conjugate Structure Algebraic Code-Exited Linear Prediction

The different coding methods vary in the levels of complexity, delay characteristics and quality. The CODEC which is expected to become prevalent within the Voice over IP arena is G.729A.

11.5. Wireless Voice over IP Networks (W-VoIP)

Both, WLAN and VoIP became central issues in communications, mainly because being very cost-effective. The ubiquity of IP networks has opened the door to reliable, high-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

quality voice applications on either wired or wireless networks. As 802.11 standard is commonly named Wi-Fi (from wireless fidelity) it becomes more and more largely used the term of Wi-Fi telephony.

Wireless VoIP encounters additional challenges based on the particularities of wireless networks:

- the need of specific QoS solutions: until 802.11e becomes an applicable standard, proprietary solution are used. The most widely known are: SVP (SpectraLink Voice Prioritization) and again SVP (Symbol Voice Prioritization).
- automatically detect and recover from interruptions: this could be cause by a dead battery or transmission loss due to impairments. In order to implement a viable solution, the existing VoiP protocols need modifications and extensions.
- higher level of security then the existing Wi-Fi technology is providing, without compromising fidelity
- interoperability with other type of wireless or wireline networks
- new challenges for high speed mobility.

Wireless VoIP can have numerous benefits:

1. for user: high mobility, convenience of a single device for multiple use, very cost-saving, possibility of tele-presence, unified management, a large variety of integrated services
2. for carriers: cost-saving delivery, increased use of the network producing additional revenue, ease of management of a unified network.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

11.6 Wireless VoIP security issues

IP based networks have two major vulnerabilities, which rely on the main advantages of this type of shared media network:

- ease of eavesdropping (taping the communication line)

- ease of spoofing (i.e. faking identity).

Since the shared media cannot be ever secure, therefore, it is necessary to secure the other components of the transmission: devices (IP-telephone), the transmitted information, the computers, servers and other network elements.

This can be accomplished by a series of tools:

- control of the access to the transmission

-authentication of both the user and the communication equipment

- encryption of the transmission

- use of call accounting package to identify the responsibility of the call and its duration.

In order to have a professional level of security of the Wireless VoIP network, some architectural adaptation of the VoIP network are necessary, like:

- to setup a strong authentication of the users and of the IP-phone

- to setup a strong authorization of the users for the various available features and services provided by the Wireless VoIP network

- to put IP PBXs in a different domain than its other servers and limit the administration access to them

- to ensure that server-based IP PBXs are protected against viruses and denial-of-service

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

attacks

- to separate and isolate voice traffic onto a virtual LAN
- to encrypt VOIP traffic by using strong encryption and multilevel encryption
- to run the voice traffic over a Virtual Private Network (VPN)
- to properly configure firewalls
- to segment voice and data traffic by using a virtual LAN; this will limit the threat posed by packet-sniffing tools and minimize disruption in the event of an attack
- to use proxy servers in front of corporate firewalls to process incoming and outgoing voice data
- to use the security tools offered by the application layer level protocol; especially SIP (Session Initiation Protocol) offer several tools as: authentication header or proxy-authentication header
- to secure the network against employee abuse, by assigning accountability, which at its turn is accomplished by distributing call accounting reports and charging back for usage

Securing the wireless VoIP network can be done at different levels of the network protocol stack.

Some of the main tools to do this, are:

1. WEP (Wireless Equivalent Privacy)

WEP is part of the 802.11 protocol, and it's used for encryption. When it is enabled, each packet is encrypted separately with a 64-bit RC4 key. This key has two parts: a 24-bit initialization vector (IV) and a 40-bit WEP key. The encryption is in fact a bitwise exclusive OR between the data packet and the key. The Initialization Vector is chosen by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the sender and can be changed. It is sent unencrypted, with each packet. An additional 4-byte Integrity Check Value (ICV) is computed on the original packet and appended to the end. The ICV is also encrypted with the key.

WEP has some important weaknesses:

- Key management is poor and key size is too small
- The IV is too small m(24 bits) and it's sent in clear
- The ICV algorithm (based on CRC-32) is not appropriate
- Authentication messages can be easily forged

In order to correct these weaknesses it is now under design the 802.11i security standard, providing authentication of both, the user and device.

It contains 2 main components:

a. Wi-Fi Protected Access (WPA) and

b. Robust Security Network (RSN).

a. WPA has the following main characteristics:

- it uses Temporal Key Integrity Protocol (TKIP) as encryption algorithm, in order to improve security of keys used with WEP.
- it changes the way keys are derived
- rotates keys more often for security
- it adds a message-integrity-check function to prevent packet forgeries.

b. RSN has the following characteristics:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- it uses dynamic negotiation of authentication and encryption algorithms between access points and mobile devices.

- The authentication schemes are based on Extensible Authentication Protocol (EAP). - The encryption algorithm is Advanced Encryption Standard (AES).

providing a much more sophisticated securing tool in Wireless VoIP networks.

2. Firewall

A firewall is a part of a security policy in the VoIP network designed to safeguard the privacy of the network.

The main functionalities of the firewall are:

- allow administrators a single point of administration and control
- filter the traffic and, hence it keeps unauthorized users (hackers or intruders) from breaking in or leaving the protected network
- simplify security management
- helps to conserve precious IP addresses

3. IPSec

IPSec protocol provide encryption and authentication tools to the network at the IP layer level. With IPSec it can be created secure tunnels through untrusted networks, which in turn can be used to create Virtual Private Networks.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

IPSec provides many options for performing network encryption and authentication. Each IPSec connection can provide encryption, integrity protection, and authentication, or all three.

In these security service different available algorithms can be used as: data-encryption standard (DES) or International Data Encryption Algorithm (IDEA) for encryption; Message Digest 5 (MD5) or Secure Hash Algorithm (SHA) for integrity.

After deciding on the algorithms, the two devices must share session keys. A security association (SA) is a relationship that is established between two VoIP entities.

In the IPSec implementation three protocols are used:

a. Encapsulating Security Payload (ESP) protocol: encapsulates user data and provides data confidentiality and protection with optional authentication and replay-detection services.

b. Authentication Header (AH) protocol: provides a packet authentication service. It is usually used in conjunction with ESP.

c. Internet Key Exchange protocol (IKE): is a hybrid protocol that uses parts of Oakley and SKEME inside the Internet Security Association and Key Management Protocol (ISAKMP) framework; It is used to establish a shared security policy and authenticated keys for services (such as IPSec) that require keys.

Before any IPSec traffic can be transmitted, each router, firewall, or host have to verify the identity of its peer. This can be done by manually entering preshared keys into both hosts, by a certification authority (CA) service, or via the forthcoming DNS secure (DNSSec).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As the security of the Wireless Voice over IP Network is a new challenging major issue, important ongoing research effort is in place and new ingenious solutions are added with every day to the canopy of already existing tools.

Bibliography

11.1. Emmerson, Bob. David Greetham. " Computer Telephony and Wireless Technologies : Future Directions in Communications". Computer Technology Research Corporation. July 1997.

11.2. Minoli, D., E. Minoli. "Delivering Voice over IP Networks". John Willey and Sons. 1998.

11.3 Minoli, D. "Delivering Voice Over IP Networks". John Willey and Sons. Oct. 2002.

11.4 Miller, M. "Voice over IP Technologies". M.&.T PUBLISHING LTD. June 2002.

11.5. Rao, H., Y. Lin, S. Chou. "iGSM: VoIP Service for Mobile Networks". I.E.E.E. Communications Magazine, 4(38), pg. 62-69. 2000.

11.6. Reynolds, J. "Going Wi-Fi: A Practical Guide to Planning & Building an 802.11 Network". CMP Books. 2003.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 12. Studying performance issues in Wireless Networks

12.1. System modeling and simulation

These day we witness an explosively diversifying and evolving communication networks.

This makes that the cost to design, build and operate these communication networks to be in a constant increase. Network modeling and simulation is an efficient way to reduce these costs, providing a cost-effective way to study the newly designed network behavior and performance, and also represent an efficient way to maintain performance parameters throughout its lifetime.

Historically, network modelers face a dilemma: to choose between speed of simulation and accuracy of results.

Simulation is an applied methodology in which one describes the behavior of complex systems, using mathematical or symbolic models.

The **model** serves as a hypothesis for how a system really behaves over time. We may also simulate non-real systems by varying parameters, initial conditions and assumptions about our model. Fictional simulations are very popular in the form of role-playing simulation games, and they permit the simulation user to learn about a new environment by interactively exploring it.

The concept of **learning** is essential in simulation. We learn about an environment in an extremely effective way and modify rules while seeing the effects of our interaction.

Computer simulation is technology, which implies the following process:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-design a model of an actual or theoretical physical system,

- execute the model on a digital computer, and

- analyze the execution output.

To understand reality and its complexity, we must build artificial objects named models and dynamically changes its parameters in different scenarios.

In doing so, computer simulation offers four big advantages:

- it's not necessary to build the model from materials;

- the parameter variation can be highly automated;

- a large number of scenarios can be effectively conceived and studied;

- even the interpretation of the results can be highly automated.

Usually simulation implies the following 3 steps:

- model design,

- model execution or

- model analysis, with two substeps:

    - verification and

    - validation.

A model is a created system used to study the behavior of a real system.

The model is not a complete low-scale reproduction of the studied system, but only a reproduction in its essential parameters. It is both art and science to determine at what level o model is a correct representation of the original system, without being a reproduction in all details of it.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

12.2. Modeling Process

12.2.1. Model Design: The first step to building a simulation model of a real system is to gather data associated with that system. Data can be in symbolic or numeric form. Usually, numerical data are obtained using either a manufactured sensor, like thermocouple, pressure meter, level meter, counter etc. or one of the five human sensors. Even if human perception could be highly subjective, using interviewing methods, brainstorming or other developed acquisition techniques valuable qualitative knowledge can be obtained.

12.2.2. Model execution: From the collected data as well as using the results of past experiments with similar systems, we formulate a model. Models could be created without data collection, but the accuracy of the model could questionable in this case. A model will start with some initial default values of its parameters. These initial values could be estimated, collected or just hypothesized. The ambiguity of a model is smaller if formal semantics is used.

12.2.3. Model Analysis:

Model analysis has two sub-steps: Verification and Validation, shortly presented below:

12.2.3.1. Models must be converted to algorithms to run on a digital computer. Verification is the process of making sure that the written computer program corresponds precisely to the mathematical model.

12.2.3.2. Validation, is the process of making sure that the model's output accurately reflects the behavioral relationships present within the original real system data. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

model, when simulated, should be able to produce the same sorts of data and input-output relationships that were initially gathered.

Next, we perform tests on the data generated from the model, by running a certain specific type of analysis, as for example, statistical analysis. The simplest type of analysis is the visual inspection of the data, followed by inferences based on this examination. Probably the most difficult problem in simulation is to determine the most efficient method to use in order to create a model.

12.3. Model Engineering

There are many modeling techniques used for simulation. Therefore a model designer has to answer two main questions:

- which model technique to use, and
- under what conditions we should use the model.

Next, the main steps to building and executing models, is presented.

1. First, a conceptual model of the system under investigation, has to be defined.
2. In a second step, several substeps could be envisioned:

2.1 the system has to be divided into a hierarchy of abstractions; this step is named multimodeling, and

2.2 the models to represent those abstraction levels, has to be chosen.

Some levels may be declarative, other could be functional, and other could be spatial in nature. Some might combine all these elements, using a hybrid modeling technique. The final solution to a model will be a multimodel of the system since no one model type will be sufficient to describe a system, except in only the most elementary circumstances.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A multimodel is a collection of individual models, each characterizing an abstraction level, and integrated together.

3. After the model has been designed, it needs to be executed. The way that one executes a model depends on the type of model. We can conceive:

-serial simulation methods

- parallel simulation methods, usually possible on special machine architectures

- distributed simulation method, which is a particular case of parallel simulation

A variety of proven modeling methods should be used, assembling them together into a multimodel.

Multimodeling is in accordance with the evolving complex, distributed and heterogeneous approach to network computing.

12.4 Models of Computation

There is a large variety of models of computation. Each gives an interaction mechanism with the components.

Further there is a short description of the main models of computation.

For a better understanding of the terminology the following definitions are given:

Any communication model can have a graph representation;

- **Entities** are nodes of the graph
- **Relations** are the arcs of the graph
- **Actors** are entities with attached functionality, and
- **Communication** between actors are the relations connecting them.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

According to the Ptolemy Project, developed at Berkeley University, we can distinguish the following type of models of computation:

12.4.1 Communicating Sequential Processes Model(CSP)

In the CSP model (communicating sequential processes) actors represent concurrently executing processes. In order for these processes to communicate, several synchronization techniques are available, the most widely used being the atomic, rendezvous concept, named also synchronous message passing. If two processes have to communicate and one reaches the point first at which it is ready to communicate, then it waits until the other process is ready to communicate too. This communication process is named atomic because the two processes are simultaneously involved in the exchange, and the exchange is initiated and completed in a single uninterruptible step.

Examples of rendezvous models include Hoare's communicating sequential processes (CSP), and Milner's calculus of communicating systems (CCS).

Examples of use are in client-server database models and multitasking or multiplexing of hardware resources. A key feature of this type of model is its ability to correctly model nondeterminate interactions.

12.4.2 Continuous Time Model (CT)

In the CT (continuous time) model of computation, also named analog model, actors represent components that interact via continuous-time signals. Actors typically specify algebraic or differential relations between inputs and outputs. The main goal in this case is to find a fixed-point (i.e., a set of continuous-time functions) that satisfy all the relations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The CT domain includes an scalable set of differential equation solvers. The domain, therefore, is useful for modeling physical systems with linear or nonlinear algebraic/differential equation descriptions.

This type of model is widely used for the study of analog circuits, microwave circuits or several type of mechanical systems, like Micro-Electro-Mechanical Systems (MEMS). Also embedded systems frequently contain components modeled using differential equations, such as and other mechanical components, analog circuits, and microwave circuits.

Sometimes these type of systems interact with digital electronic systems.

The joint modeling of a continuous subsystem with digital system is known as mixed signal modeling.

12.4.3 Discrete-Events Model (DE)

In the discrete-event (DE) model of computation the actors communicate via sequences of events placed in time, along a real time line. An event consists of a value and a time stamp.

Actors can either be processes that react to events or functions that execute when new events are supplied. This model of computation is popular for specifying digital systems, like telecommunications systems, and has been realized in a large number of simulation environments, simulation languages, and hardware description languages .

Every event is placed precisely on a globally consistent time line.

DE simulators in general need to maintain a global queue, named priority queue, containing pending events sorted by time stamp. Every event is placed precisely on a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

globally consistent time line. Sometimes, in very fast changing systems this could be both difficult and expensive, both the insertion and the searching in the queue.

The DE domain uses a calendar queue data structure for the global event queue. A calendar queue may be conceived as a hash-table that uses quantized time as a hashing function.

Both, enqueue and dequeue operations can be done independently, from the number of events in the queue.

The DE domain gives deterministic semantics to simultaneous events, as opposed to other discrete-event simulators. This means that for any two events with the same time stamp, the order in which they are processed can be inferred from the structure of the model, by choosing a secondary criterion to define precedence relationship.

12.4.4 Distributed Discrete Events Model (DDE)

The distributed discrete-event (DDE) model of computation can be viewed either as a variant of DE or as a variant of PN (described below). The DE model has the global event queue, which imposes a central point of control on the model, and therefore greatly limits the ability to distribute a model over a network. DDE is solving this limitation, which could be very useful or even necessary in order to conserve network bandwidth, or to exploit parallel computing resources.

The DDE domain maintains a local notion of time on each connection between actors, instead of a single globally consistent notion of time. Each actor is a process, that can advance its local time to the minimum of the local times on each of its input connections.

12.4.5 Discrete Time Model (DT)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The discrete-time (DT) model of computation extends the SDF domain (described below) with a notion of time between tokens. Communication between actors takes the form of a sequence of tokens, where the time between tokens is uniform.

A particular type of DT models is the Multirate Model, where distinct connections have distinct time intervals between tokens. The semantics is defined in such a way that every component behavior is always causal, which means that an output whose value depend on the input is never produced at time prior the depending input.

12.4.6 Finite-State Machines Model (FSM)

The finite-state machine (FSM) model of computation has the entities represented not actors but rather **state**, and the connections represent transitions between states.

Execution is a strictly ordered sequence of state transitions.

FSM models are suitable to in-depth formal analysis, and thus can be used to avoid surprising behavior.

The main advantage of FSM model is that many questions about design are decidable for FSMs and undecidable for other models of computation.

FSM models have some weaknesses, as:

- they are not as expressive as some other models of computation
- they are not sufficiently rich to describe all partial recursive functions.
- the number of states can get very large even in the face of only medium level of complexity.

However, this weakness is compensated by the formal analysis, which becomes possible.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The drawback of FSM models can be compensated by using them in combination with other models, mainly CM (Concurrent Models of computation). When combined with CT, it result either hybrid systems and/or modal models.

When combined with SR (Sequential/Reactive, described below), they resemble to statecharts.

When combined with PN (Process Networks, described below), they resemble to SDL.

FSM models are widely used and are excellent for control logic in embedded systems, as well as for safety-critical systems.

12.4.7 Process Networks Model (PN)

In the process networks (PN) model of computation processes communicate by sending messages through channels that can buffer these messages. The sender of the message need not wait for the receiver to be ready to receive the message. This communication process is named asynchronous message passing.

There are several variants of this modeling technique, the most widely known being Kahn process networks.

In the PN model of computation, the arcs represent sequences of data values (tokens), and the entities represent functions that map input sequences into output sequences. Certain technical restrictions on these functions are necessary to ensure determinacy, as for example the function implemented by an entity must be prefix monotonic. The PN domain realizes a subclass of such functions, first described by Kahn and MacQueen, where blocking reads ensure monotonicity.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

PN models are loosely coupled, and therefore relatively easy to parallelize or distribute. They can be implemented efficiently in both software and hardware, and leave implementation options open.

An important weakness of PN models is that they are awkward for specifying control logic, but it can be eliminated by combining them with FSM models.

12.4.8 Synchronous Dataflow Model (SDF)

The synchronous dataflow (SDF) model of computation handles regular computations that operate on streams. Dataflow models, used in signal processing, are a special case of process networks. Dataflow models represent processes of a process network, as sequences of atomic actors. Synchronous dataflow (SDF) is a special case of Dataflow model, in which deadlock and boundedness are both decidable, property that makes it very useful in some cases. Also, the sequencing of actors can be scheduled in parallel or in serial, and is statically computable, making it an useful specification formalism for embedded real-time systems.

Boolean dataflow (BDF) models sometimes yield to deadlock and boundedness analysis, although fundamentally these questions are undecidable.

Dynamic dataflow (DDF) uses only run-time analysis, and thus makes no attempt to statically answer questions about deadlock and boundedness.

Process networks (PN) can handle computations that do not match the restrictions of SDF.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

12.4.9 Synchronous/Reactive Model (SR)

In the synchronous/reactive (SR) model of computation, the arcs represent data values that are aligned with global clock ticks. Thus, they are discrete signals, but unlike discrete time, a signal need not have a value at every clock tick. The entities represent relations between input and output values at each tick, and are usually partial functions with certain technical restrictions to ensure determinacy.

Because of the tight synchronization, SR models are excellent for applications with concurrent and complex control logic. like real-time applications. But for the same reason, some applications are over-specified in the SR model, thus limiting the implementation alternatives. Also in many realization of the model, modularity is compromised by the need to seek a global fixed point at each clock tick.

The above presented models of computation can be combined in two different ways:

1. Modal Model: is one that switches between these simple models when the system transitions between regimes. For example the CT domain interoperates with the FSM domain to create modal models.

2. Hybrid Model: is one that contains in the same basic model, characteristics from two or more different models of computation For example the same model could be a finite state machine where some parts are computed with differential equations, so the same basic model contains both FSM and CT.

12.5. Network Simulators

There is a large number of Network Simulation packages available either as freeware, shareware or as commercial products.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Below we present a list of the most remarkable ones, with a short description of their main characteristics and functionalities:

12.5.1 OMNeT++ (Objective Modular Network Testbed in C++), is a discrete event simulation tool developed by Andras Varga from Hungary, and occasional contributors from all over the world. It was primarily designed to simulate computer networks, multi-processors and other distributed systems, but it may be useful for modeling other systems as well. In the past few years, OMNeT++ become a popular network simulation tool in the academic and scientific community as well as in some industrial environments.

The main components of OMNeT++ are:

- simulation kernel library
- compiler for the NED topology description language (nedc)
- graphical network editor for NED files (GNED)
- GUI for simulation execution, links into simulation executable (Tkenv)
- command-line user interface for simulation execution (Cmdenv)
- graphical output vector plotting tool (Plove)
- utilities (random number seed generation tool, makefile creation tool, etc.)
- documentation, sample simulations, contributed material, etc.

OMNeT++ has been developed on Linux, but it works on most Unix systems and on Microsoft Win32 systems.

As a freeware, OMNeT++ has its source code freely available, which makes it appealing for further development

The main drawback is the relatively small number of available simulation models.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

12.5.2 Ptolemy: The Ptolemy system is mainly used for the modeling, simulation, and design of concurrent, real-time, embedded systems. It is developed in the Department of Electrical Engineering and Computer Sciences of the University of California at Berkeley, under the leadership of Professor Edward Lee. The focus is on systems that mix technologies, as for example analog and digital electronics, hardware and software, and electronics and mechanical devices, and also on systems that are complex in the sense that they mix widely different operations, such as signal processing, feedback control, sequential decision making, and user interfaces.

Ptolemy is an extensible, open architecture system, publicly available. It runs on most Unix type systems. Originally it was written in C++. An improved version named Ptolemy II, is written in Java.

It is widely used in both academic and industrial environment

12.5.3 cnet network simulator enables experimentation with various data-link layer, network layer, routing and transport layer networking protocols in networks consisting of any combination of point-to-point links and IEEE 802.3 Ethernet segments. cnet was written and is maintained by Chris McDonald from University of West Australia. **cnet** has been specifically developed for, and is used in, undergraduate computer networking courses taken by thousands of students worldwide since 1991. **cnet** runs on a variety of UNIX and Linux platforms. It does not run on either Windows or the Apple Macintosh. cnet either displays the entire network under Tcl/Tk or runs rather less visually on an ASCII terminal. Under Tcl/Tk, cnet provides a graphical representation of the network

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

under execution and permits a number of attributes of the network to be modified while the simulation is running. Nodes may be selected with the mouse to reveal a sub-window displaying the output and protocol statistics of that node. Some of the node's attributes, such as message generation rates and sizes, may be modified while the network is running by selecting choice buttons. Similarly, the default attributes of all nodes in the network may be simultaneously modified by selecting and changing global attributes. From another menu, each node may be forced to reboot, (impolitely) crash, (politely) shutdown and reboot, pause and (hardware) fail.

cnet requires network protocols to be written in the ANSI-C programming language and supports their execution within a single UNIX or Linux process. A standard compiler, preferably gcc, is used to compile the user-written protocol code. The compiled code is then dynamically linked at run-time with the cnet simulator. Protocols are invoked directly by cnet itself - they are not interpreted. By design, the protocols do not need to contain any windowing code.

12.5.4 ns is an other widely used network simulator distributed as a shareware. It began under the name of "REAL network simulator" in 1989 at University of California at Berkeley, and has evolved substantially over the past few years. From 1995 ns development was supported by DARPA, through the VINT (Virtual InterNetwork Testbed) project at LBL, Xerox PARC, UCB, and USC/ISI. Ns has also included important contributions from other researchers, including wireless code from the UCB Daedelus and CMU Monarch projects, and Sun Microsystems.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Ns is a discrete event simulator targeting the research for a large variety of at networking architectures and topologies for both wireline and wireless connections.

ns is written in C++ and it runs on virtually any UNIX like systems, under MS Windows and MAC. New functionalities are added and a restructuring of the original design is made under the name of ns-2.

For simulation in ns, two programming languages are used: C++ is used for protocol implementation and OTcl is used for simulation configuration. TclCL is the glue library that facilitates the share of functions and variables.

Ns is made of a core event scheduler and a large library of implemented protocols, and models.

A large variety of network protocols on each layer level are already implemented and tested.

Real-time simulation and visualization can be made with the companion product: nam (network animation module). Nam also has an editor capable to generate ns scripts. Nam supports topology layout, packet level animation, and various data inspection tools.

Post-run analysis can be made using other scripts written in awk, Perl or Tcl.

12.5.5 SSFNet (Scalable Simulation Framework for Networks) is an other public domain simulation and modeling software, initiated by DARPA's ITO project S3 and today maintained by Renesys Corporation.

It is a discrete event simulations package mainly conceived for large, complex network systems, with available Java and C++ high-performance scalable parallel implementations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

SSFNet models are self-configuring, which means that each SSFNet class instance can autonomously configure itself by querying a configuration database, which may be locally resident or available over the Web. The network configuration files are in the Domain Modeling Language format. They are used to configure a complete model, and create a simulation with the help of the scalable DML configuration database package that is distributed with the SSF simulators.

There are manly two derivative frameworks, which can be used to create virtually any Internet model:

- SSF.OS , used for the modeling of the host and operating system components, including protocols, and

- SSF.Net, used for the modeling of network connectivity, or for creating node and link configurations.

The frameworks SSF.OS and SSF.Net hide all the details of the discrete event simulator SSF API, giving the possibility to protocols in a similar way to a real operating system

12.5.6 QualNet is a family of tools used for network design, simulation and planning, designed and maintained by Scalable Network Technologies. It can run on most UNIX or Windows platforms.

The main component parts of QualNet are:

- QualNet Developer is used to build and execute custom models or to enhance existing models from the QualNet Library of Models. Additionally, it provides debugging and analysis tools in order to speed up the development process.

It contains the following modules: animator, designer, analyzer, tracer and simulator.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- QualNet Parallel Developer is used to build and execute models on parallel computers for larger or faster simulations.

- QualNet Model Library provide network designers with a large library of networking applications. The QualNet model library consists of : Standard Library, MANET Library, QoS Library, Satellite Library and Cellular Library.

- QualNet Integration Modules gives the possibility to link to other simulations and technologies. The main modules are:

    - importer, with topology and traffic importing capabilities,

    - terrain module, mainly use din wireless simulations, and a

- high level architecture and threaded communications module, mainly used in warfare network simulations.

12.5.7 GloMoSim (Global Mobile Simulation Library) is a scalable simulation environment, conceived and maintained at UCLA. It's a discrete event simulator mainly used for wireless network systems. In the future it is anticipated to provide simulation tools also for wired networks and hybrid wireless-wired networks. It can be used only in association with Parsec. Parsec is a simulation language, conceived as an extension of the "C" language, for sequential and parallel execution of discrete-event simulation models. It can also be used as a parallel programming language. GloMoSim runs on both UNIX or Windows platforms.

Similar to the OSI seven layer network architecture, GloMoSim is using a similar layered approach. Standard APIs are used between the different simulation layers. This gives the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

possibility for a rapid integration of models developed at different layers by different people.

Also it has a Visualization Tool, written in Java, so platform independent, which gives the possibility to see real-time evolution of a network and to gather and see statistical data.

12.5.8 COMNET III is a commercially available design, simulation and performance analysis tool for communications networks. It is used to model networks, their control algorithms, and workload. COMNET III then simulates the operation of the network and provides measures of network performance. COMNET was produced initially by CACI Corporation from La Jolla, CA and is now maintained by Compuware Corporation from Detroit, MI. It runs on most of UNIX and Windows NT based systems.

The main components of COMNET are: Network Topology, Network Traffic and Workload, Network Operation, Simulation Control, Statistics Reporting, User Distributions, Libraries, and Model Files. It contains an extensive library of components as well as pre-existing models.

COMNET III has an object-oriented design and it allows the simulation of complex network systems, being highly scalable. It has several add-on modules, which can be used to simulate circuit-switched, cellular or satellite communications networks.

12.5.9 ONE (the Ohio Network Emulator) is a freeware network simulation package produced at Ohio University Internetworking Research group and is now maintained in conjunction with the NASA Glenn Research Center's Satellite Networks and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Architectures Branch. The package works on under Sun's SOLARIS operating system. The user can control various features of the network, such as propagation delay, queueing characteristics and bandwidth. In addition, the tool interfaces with Satellite ToolKit (STK) to emulate variable propagation delays based on the orbits of satellites.

12.5.10 OPNET Modeler (OPtimized Network Engineering Tools) is probably the most elaborated, commercially available communications network simulation and design tool. It was initially produced at MIT around 1987, and later became a commercial product, maintained and developed by mil3, a company renamed now OPNET, from Bethesda, MD.

OPNET can simulate almost any communication network architecture, protocols or communication links. OPNET has an open systems methodology and an advanced graphical user interface for building, simulating, and analyzing network models.

Since in the wireless network performance analysis of this work, we will use OPNET, a more detailed description of the package follows.

The main components of OPNET are:

- graphical representation of models, ;
- dynamic simulation kernel;
- integrated data analysis tools;
- statistical distribution functions and;
- hierarchical object based modeling.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

OPNET's hierarchical modeling structure give the possibility to simulate complex architectures and to solve special problems as, for example, distributed algorithm development.

OPNET analyzes the system behavior and its performance by using discrete event simulation techniques that model system behavior around objects and distinct events. In OPNET, a network is designed by using different objects, which are the elements of a network and have several attributes to characterize them. For example, transmission link is a network element with attributes like transmission rate, link type (wire, optical fiber, radio or satellite), link length, etc. During the simulation process, the attributes associated with various objects control the behavior of those objects. OPNET Modeler is associated with several support of interrelated programs, model libraries, and data files.

The primary tool of this system is the OPNET Modeler program. The key features include object orientation, graphical specification, automated model creation, an extensive model suite of various manufacturer's devices, integrated analysis tools and animation support.

The model specification can be developed in the OPNET using a large range of editors. Editors used in the OPNET to develop the specifications are:

- Project editor: is used to develop a network model. A network model consists of subnets, nodes, protocols/processes and transmission links. This editor also has basic simulation and

analysis capabilities.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Node editor, is used to develop node models. Node models are objects in a network model. Node models are made up of modules with process models that describe various operating and functional capabilities.

- Process editor is used to develop process models. Process models control module behavior and may have many reference parameter. For example a process model within a PC could represent Ethernet transmission and reception of packets.

- Link editor is used to create, edit and view transmission link models.

- Packet editor is used to develop packet format models. Packet format dictate the structure and order of information in a packet.

- PDF (Probability Density Function) editor is used to create, edit and view probability density functions. PDF's can be used to control certain events such as frequency of packet generation in a source model.

- Probe editor is used to specify what simulation output data should be collected.

- Analysis editor is used to analyze simulation result. Advanced statistical analysis functions are available.

12.6. Simulation Models, using OPNET

In OPNET, a network model consists of a number of nodes (or terminals) and communication links. Information flows from node to node using communication links, and the transmitted bits are grouped in packets.

- The packet format defines the structure of control fields and of the information field. For different networks we'll have different packet format.

- Link model contains all the parameters of communication links used. The main link parameters are: link type (simplex, duplex), transmission rate, bus/tree configuration,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

propagation delay, error model (to simulate bit error), etc. There are a number of standard link models already in the OPNET library but also custom made links can be created.

- A node model consists of a number of processes which basically describes the operating behavior of the node. Node models have a hierarchical structure where, on top level shows modules (blocks) and underneath those modules are process models, which describe the behavior of each module.

- Process models are coded using the OPNET commands and kernels. There are a large number of standard node models available in the standard library as well as in several third party vendor libraries.

To develop a network model the user need to integrate the nodes and links. After creating the network model, the behavior of the model can be analyzed by defining appropriate statistics. The results of the analysis can be visualized in text, table or graphic format.

Bibliography

12.1. Fishwick, Paul. "Computer Simulation". At: http://www.cise.ufl.edu/~fishwick/introsim/paper.html.

12.2. Heidemann, John, N. Bulusu, J.Elson, et al. "Effects of Detail in Wireless Network Simulation". Distributed Systems Modeling and Simulation Conference. Jan. 2001.

12.3. Heidemann, John, K. Mills, S. Kumar. "Expanding Confidence in Network Simulations". IEEE Network. Sept/Oct. 2001.

12.4. http://ptolemy.eecs.berkeley.edu/ptolemyII/ptII3.0/ptII3.0/doc/

12.5. http://www.omnetpp.org

12.6. http://ptolemy.berkeley.edu/

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

12.7. http://www.isi.edu/nsnam/ns/

12.8. http://www.isi.edu/nsnam/ns/tutorial/

12.9. http://www.ssfnet.org/internetPage.html#ssfnet

12.10. http://www.scalable-networks.com/products/qualnet.php

12.11. http://deim.etse.urv.es/ajuda/manuals/comnet/c3ref.pdf

12.12. http://www.managementsoftware.hp.com/sso/isv/detail?appid=A668

12.13. http://jan.ucc.nau.edu/~ejacobs/Comnet.htm

12.14. http://www.opnet.com/products/modeler/home.html.

12.15. Johnson, D. B. "Validation of Wireless and Mobile Network Models and Simulation". Proceedings DARPA/NIST Network Simulation Validation Workshop. Fairfax, VA. May 1999.

12.16. Schwartz M. "Telecommunications Networks: Protocols, Modeling and Analysis". Addison-Wesley. Reading, MA. 1987

12.17. Zeigler, Bernard P., H. Praehofer, T.G. Kim. "Theory of Modeling and Simulation", Second Edition. Academic Press, San Diego. 2000.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 13. Study of Performance Issues in Wireless Multimedia Networks

13.1. Wireless network modeling, using OPNET simulator

As we intend to study performance issues in Wireless networks using the OPNET simulation package, further, a detailed presentation of WLAN modeling using OPNET, is presented.

Some of the important features of the WLAN model using OPNET simulator are as follows:

1. **Wireless LAN Protocol Behavior Modeling**

As Access Mechanism was implemented Carrier sense multiple access and collision avoidance (CSMA/CA) DCF access scheme according to the standard.

The following Frame exchange sequence are modeled:

- Data and Acknowledgment frame exchange, to ensure the reliability of data transfer.
- Optional RTS/CTS frame exchange, for media reservation.

2. Deference & Backoff

It was implemented:

- DIFS, SIFS, and EIFS interframe spacing, for DCF. The values for the intervals are selected based on the selection of physical characteristics.
- Binary exponential backoff.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. Data Rate The following data rates supported by the WLAN protocol are implemented: 1Mbps, 2 Mbps, 5.5 Mbps, and 11 Mbps. We will make use only of the 11Mbps transmission rate.

4. Recovery mechanisms

As recovery mechanism, a retransmission mechanism for data frame, based on the failure of the reception of the acknowledgment frame, was implemented. It is based on short and long retry counters, as defined in the standard.

5. Fragmentation and Reassembly

Optional Data frame fragmentation is implemented, based on the size of the data packet received from the higher layer. The fragments are reassembled at the destination station.

6. Duplicate Packet Detection

In order to detect duplicate packet arrival, a cache mechanism is implemented, in order to store the information of the received packet. As a result, any duplicate packets are discarded by the MAC layer.

7. Access Point Functionality

The implementation allows that any station to be able to be configured as an access point in an infrastructure BSS network. Always, currently only WLAN routers have a capability of connecting a BSS to a network distribution system. In order to form an ESS, allowing stations from different BSSs to communicate with each other, an IP network needs to be configured.

Note: Each OPNET Subnet is considered to be one BSS.

8. Message Buffer Size

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Data arrived from a higher layer to the WLAN MAC is stored in a buffer. This buffer size is limited to the maximum value set by the user. Higher layer packets are dropped, once the maximum buffer size is reached.

9. Radio IP Autoaddressing

All the WLAN nodes can be configured for IP autoaddressing. The nodes within one OPNET subnetwork are given the same network address. If there is any static assignment then that will be the network address of all the nodes within the subnet. However, if there are more than one static assignment with different network addresses then the first encountered static assignment will be the network address of all the nodes in the subnet. Note: Wireline nodes have a separate procedure for auto-addressing, different the radio IP auto-addressing.

10. Physical Layer

The physical layer dependent parameters (like timing intervals and Backoff slots) are modeled in the WLAN MAC.

The model simulates up to a certain level, the actual physical layer IEEE 802.11specification. There are some simplifications made to the radio pipeline stages, as follows:

- Radio receiver: The receiver group of the station is restricted to its subnet and the transmitting station will not receive its transmitted packets;

- Channel match: The transmitter and receiver data rate should match in order to successfully transmit a packet

- Propagation delay: There is a maximum range across which the station can communicate. This pipeline stage will discard the packet if the destination is beyond the specified range;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Error correction: If the receiving packet contains errors bigger than the error threshold, then the pipeline stage will mark it as a corrupted packet and the MAC layer will discard it.

13.2. Elements of the OPNET WLAN simulation model

Further the key elements of the WLAN simulation model, including model attributes and statistics, are presented. The main goal is to show how the model simulates the protocol behavior, but some of the implementation specific details are not presented.

As mentioned before, there are two types of wireless network configuration: Infrastructure BSS and Independent BSS. The current WLAN model supports both the configurations and their setup is as follows:

1. Infrastructure BSS

Workstations can communicate outside of the BSS using Access Point (AP) which is connected to the distribution system. At the same time, this AP acts as a relay within the BSS.

2. Independent BSS

Workstations cannot communicate outside of the BSS.

WLAN Object Palette

All the project scenarios can be created using models available as part of WLAN object palette.

The main models, included in the OPNET library, are:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. Wireless Workstation (fixed and mobile)

2. Wireless Server (fixed and server)

3. Wireless terminal station (WLAN MAC without IP)

4. Wireless Router (with one wireless interface and one or more wireline interface like Ethernet, FDDI, Token Ring, ATM etc.)

Further we shortly describe each of these node models:

1. Wireless Workstation

The model has implemented at physical and MAC level, the WLAN MAC process, transmitter, receiver, and the channel streams. The ARP (Address Resolution Protocol) is an interface between the MAC and the higher layers. The AP attribute or a Wireless Workstation toggles the on/off switch of the station's access point functionality.

2. Access Point and Wireless Router

The access point or router, contains one wireless interface and one Ethernet interface. It can be used to connect a BSS with a wireline network distribution system. The wireless interface receives the data frame and sends it to the higher layer for address resolution, if the destination station does not belong to the same BSS. There can be only one AP in a BSS. If the user tries to configure more than one AP in a BSS, then the simulation will abort with an error message, indicating that multiple APs are configured in BSS.

3. Basic Service Set (BSS)

WLAN is configured by using OPNET subnet object. Each OPNET subnet can be considered as a single BSS if the user intends to auto configure the IP addresses. Also, this

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

object is used to represent logical and physical grouping of nodes within a larger topology. Once all

the subnet are appropriately configured then they can be connected through a distribution system, using for example a wireline technology.

WLAN Parameters

Further the WLAN-related configurable parameters available as part of the simulation model, are presented. There are two types of attributes:

- model attributes, which have local significance on nodes, and
- simulation attributes, which are applicable across all nodes in the network.

Each wireless node has the same set of wireless LAN attributes. These attributes are grouped under Wireless LAN Parameters. There is also a Wireless LAN MAC Address parameter which is an internal station address and is usually set to "auto assigned", unless specific configuration is required..

The various WLAN attributes and their effect on the WLAN MAC are explained next:

- RTS Threshold: Specifies a threshold to determine whether or not Rts/Cts (request to send/Clear to send) frame exchange is required for a particular data frame. If the MAC service data unit (MSDU), received from higher layers in the protocol stack has a size greater than the Rts threshold, then Rts/Cts exchange is needed for media reservation. The default value for this attribute is NONE, which means that no Rts/Cts exchange will take place regardless of the MSDU size. Note that Rts/Cts exchange is optional in the protocol, but all stations should be able to respond to the Cts sent by the remote station.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

When Rts/Cts frame exchange is enabled, then there is an additional overhead involved in the successful transmission of the frame. Although, Rts/Cts exchange improves the reliability of data transmission, the channel reservation time per data frame might increase.

- Fragmentation Threshold Specifies a threshold to decide if a MSDU received from the higher layer needs fragmentation before transmission. The number of fragments to be transmitted is calculated based on the MSDU size and the fragmentation threshold. The default value for this attribute is NONE, which means that fragmentation will not take place regardless of the MSDU size. The destination station will receive these fragments and store them in the reassembly buffer until all fragments have been received. This fragmentation and reassembly is implemented using

the built-in SAR (segmentation and reassembly) package in OPNET. Using fragmentation for larger sized packets improves the reliability of data exchange between the stations. Since every data fragment requires an acknowledgement, the overall frame exchange per MSDU is higher than it would be without fragmentation. Fragmentation may also impact the total channel

reservation time per data frame exchange.

- Data Rate The WLAN model supports data transfer at 1 Mbps, 2 Mbps, 5.5 Mbps, and 11Mbps. These data rates are modeled as the speed of the transmitter and receiver connected to the WLAN MAC process. Each data rate is associated with a separate channel stream, from the MAC process to the transmitter and from the receiver to the MAC process. A station can only transmit data packets at the data rate specified by the attribute. However, it can receive data at any data rate. All control packets are transmitted at a data rate of 1 Mbps, as specified by the standard.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Physical Layer: There are three choices for the physical layer configuration in IEEE 802.11 standard: Frequency Hopping, Direct Sequence and Infra Red. Physical layer specifications are not explicitly modeled. However, physical layer parameters needed by the MAC are configured based on user selection of the physical layer. The parameters which are set based on the physical characteristics are:

1. Short Interframe Spacing (SIFS)
2. DCF Interframe Spacing (DIFS)
3. Minimum and Maximum contention window size, necessary to select number of backoff slots.

The default setting for Physical Layer is Frequency Hopping.

- Short Retry Limit This attribute indicates the maximum number of transmission attempts before a data frame is discarded. This retry limit is applicable to a data frame whose MSDU size is less than or equal to the Rts threshold (i.e. the frame for which Rts/Cts exchange is not needed). The default value is set to 7 attempts.

- Long Retry Limit This attribute indicates the maximum number of transmission attempts before a data frame is discarded. This retry limit is applicable to a data frame whose MSDU size is greater than or equal to Rts threshold (i.e. the frame for which Rts/Cts exchange is needed). The default value is set to 4 attempts.

- Channel Settings In the WLAN model there are four transmitting and receiving channels. Each channel is associated with a data rate of 1, 2, 5.5, and 11 Mbps. The user can set the minimum frequency and the bandwidth of these channels. These settings will be for all four channels since they belong to the same station. At a certain moment only one channel is in use.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Buffer Size This attribute specifies the maximum length of the higher layer data arrival buffer. Once the buffer limit is reached, the data received from the higher layer will be discarded until some packets are removed from the buffer.

- Max Receive Lifetime: This attribute indicates the maximum time after which further attempts to reassemble the MSDU at the destination station will be terminated. The timer starts after the initial reception of the fragmented MSDU.

- Simulation Attribute: The WLAN model defines one simulation attribute, which can be modified from the configuration simulation panel.

- Wireless LAN Range: This simulation attribute specifies the operating range of the stations in the BSS. Stations beyond that range cannot communicate with each other. The maximum allowable range is 300 meters which is derived from a 1 microsecond air propagation time between the stations, as specified by the IEEE 802.11b standard.

WLAN Statistics

In order to analyze the performance of the WLAN protocol, several statistics can be collected during simulation execution. Statistics can be collected on a per-node or a per-module basis.

I. Node Statistics

The implemented node statistics are:

1. Load: represents the total number of bits received from the higher layer. Packets arriving from the higher layer are stored in the higher layer queue.

2. Media Access Delay: is the total time (in seconds) the packet is in the higher layer queue from arrival, to the point when it is removed from the queue, for transmission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. Throughput: is the total number of bits sent to the higher layer from the MAC layer. The data packets received at the physical layer are sent to the higher layer if they are destined for this station.

II. Module Statistics

Module statistics are collected on a per station basis. The main implemented module statistics are:

1. Backoff Slots
2. Channel Reservation (NAV counter)
3. Control Traffic Sent (includes Ack, Rts, and Cts)
4. Control Traffic Received (includes Ack, Rts, and Cts)
5. Data Traffic Sent
6. Data Traffic Received
7. Retransmission attempts
8. Dropped data packets (due to the overflow of higher layer buffer)

13.3. Simulations for the Study of Performance Issues in Wireless VoIP Networks

We model several architectures of multimedia wireless networks and through simulations, we study the impact of different parameters, on the quality of voice transmission over the network.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

All simulations are made using OPNET simulator with Radio interface, running on a PC. The configuration of the PC is:

- AMD - Athlon microprocessor at 1000MHz, overclocked to 1200MHz
- 512 Mb internal memory, with 100MHz access speed, overclocked to 116MHz
- 30 GB hard disk with 100 Mbps access speed
- Windows 2000 Professional operating system
- OPNET 8.0c simulator with Radio interface.

Three types of wireless network architectures are considered for voice transmission:

- Ad-hoc wireless LAN,
- Infrastructure (named also backbone) wireless LAN
- Wireless LANs, interconnected with wireline routers.
- In all scenarios, 802.11 WIFI wireless communication protocol is simulated at 11Mbps transmission speed.
- UDP over IP is used as transport layer level protocol.
- As multimedia communication, digitized voice over packet switching network is generated.

Several variations of the multimedia communication are modeled and simulated, and the performance of the network is studied in the following situations:

1. First we consider VoIP users in an ad-hoc wireless network and we study if the essential voice quality parameters: delay, jitter and network saturation are reached.

The quality of the voice transmission is considered acceptable if the parameters are below the following thresholds:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- delay is less than 250 milliseconds (ms)

- delay variation (jitter) is less than 100 milliseconds (ms)

- packet drop rate is bellow 16bps.

In this network we made the following performance studies:

- we compare which physical transmission is more suitable for voice communication: Frequency Hopping or Direct Sequencing;

- we compare which voice encoding scheme is more efficient: G.711 or G.729;

- we study the influence of the mobility on performance, by comparing a network with 32 fixed wireless stations with a network with 32 mobile wireless stations

2. Further, using the best physical transmission scheme, the best voice codec, we studied what is the impact of the number of simultaneous calls, able to use the wireless network in the limits of voice quality parameters: delay, jitter, congestion;

The following voice quality parameters are analyzed, in numerical and graphical representation:

- Voice packet delay variation (jitter)

- Voice packet end-to end delay

- Voice traffic sent (bytes/sec)

- Voice traffic received (bytes/sec)

- Dropped packets (pkt/sec)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The performance parameters are analyzed at global network level or at individual wireless workstation level.

Below there is a snapshot of the general configuration of the network:

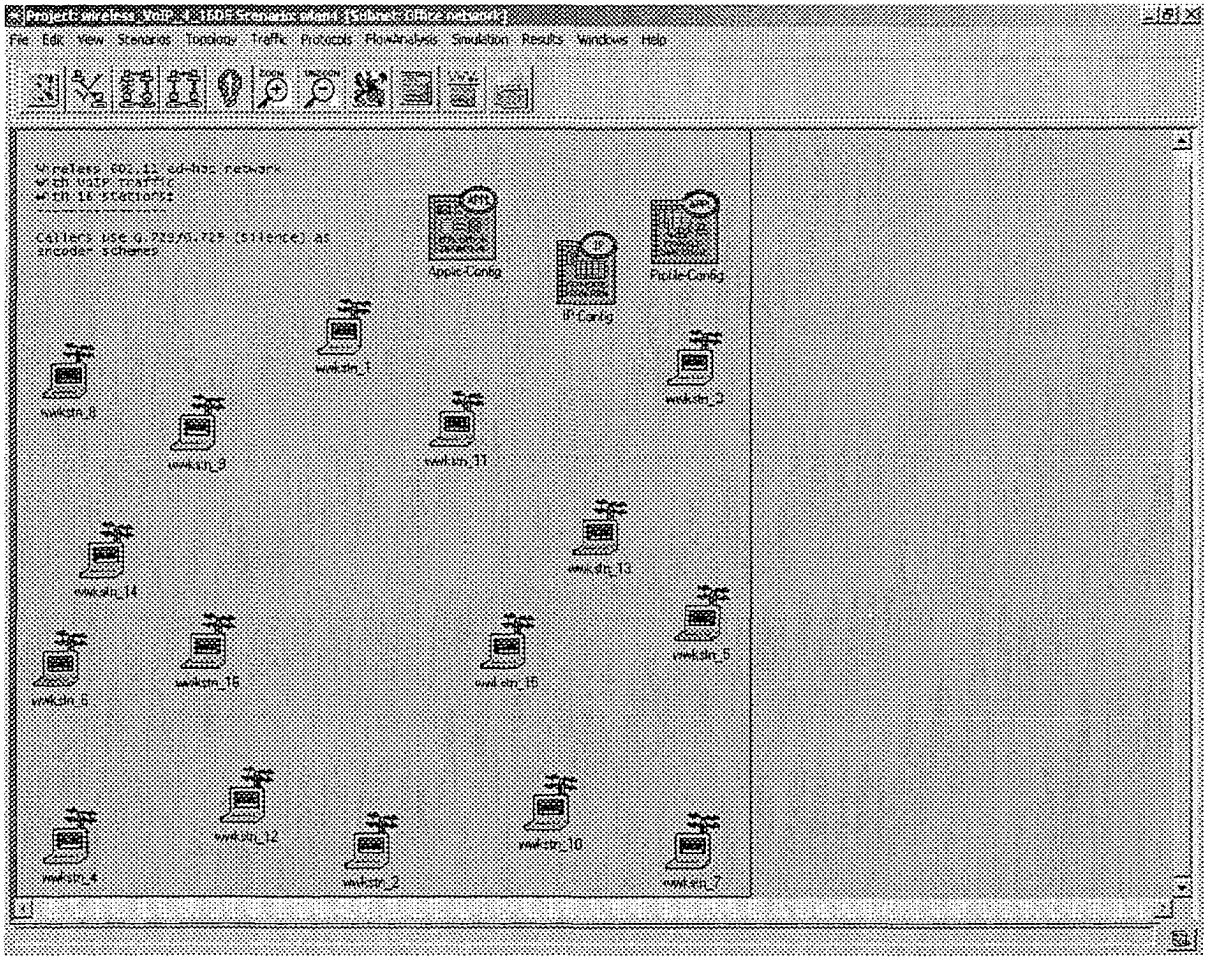


Figure 13.0. General Architecture of a simulated WLAN

In order to create the simulation model, 4 types of configurations are defined:

- Application Configuration

- Profile Configuration

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- IP protocol Configuration

- QoS Configuration

Application Configuration is used for the following specifications:

Application Definition: specifies applications using available application types. It is used to specify a name and the corresponding description in the process of creating new applications.

In our case, we use Application Definition to specify the application "Voice-IP Telephony", with the following attributes and corresponding values:

1. Silence Length, with two sub-attributes:

- Incoming Silence length: determines the time spent in silence mode, in seconds, by the called party. It is set as a random variable with exponential distribution of mean outcome value of 0.65.

- Outgoing Silence Length: determines the time spent in silence mode, in seconds, by the calling party. It is set as a random variable with exponential distribution of mean outcome value of 0.65.

2. Talk Spurt Length: which specifies the time spent in talk mode by the two parties, in a talk-silence cycle. It has two sub-attributes:

- Incoming Talk Spurt Length: specifies the time spent in talk mode, in seconds, by the called party. It is set as a random variable with exponential distribution of mean outcome value of 0.352

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Outgoing Talk Spurt Length: specifies the time spent in talk mode, in seconds, by the calling party. It is set as a random variable with exponential distribution of mean outcome value of 0.352

3. Symbolic Destination Client: specifies the destination client for the call originating from this node. It is set to the symbolic value "Voice Destination".

4. Encoder Scheme: specifies the encoding scheme used by both the calling and the called party. In all simulations we will use g.711.or g.729 as encoding schemes.

5. Voice Frames per packet: specifies the number of encoded voice frames grouped in a voice packet sent by the application to the lower layers. We will use in all simulation the value of 1.

6. Type of Service: specifies the quality of service of each IP packet sent by the client. It simulates the TOS field of the IP protocol. Any value from 0 to 7 is accepted. The higher value gives higher priority in the delivery of the packet. We will compare the performance in case of the lowest and highest priority. More precisely the value "0" is attached to the ToS named "Best effort" and the value "6" is attached to the ToS named "Interactive Voice".

7. Voice Encoder Schemes: specifies encoder parameters for each of the encoder schemes used for generating voice traffic in the network. We will compare the performance of the network in case of using the g.711 and g.729 encoding schemes with silence.

The encoding schemes with silence have the following sub-attributes set:

- Frame size: specifies the processing time in seconds of the incoming voice frame, and is set to 0.004 sec.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Lookahead size: is the amount of time in seconds, the frame is delayed for compression and lookahead calculations. This parameter is not used and the value is set to "0".

- DSP processing ratio: specifies the ratio between the DSP processing time and the frame size in seconds. Normally the value should be greater then 1.0. We will consider the ideal value of "1".

- Coding rate: specifies the amount of bits processed in a second. For g.711 it is set to 64kbits/sec and for g.729 it is set to 8kbits/sec.

- Speech activity: specifies the control of the silence compression. If it is set to "Enabled" then during silence there will be no data sent over the network. We will use in all cases the attribute set to "Enabled", which means that during the randomly generated silence periods will be no data sent over the network.

Note: We do not use the Application node for "ACE Tier Information".

For details about the setting see the figures below.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

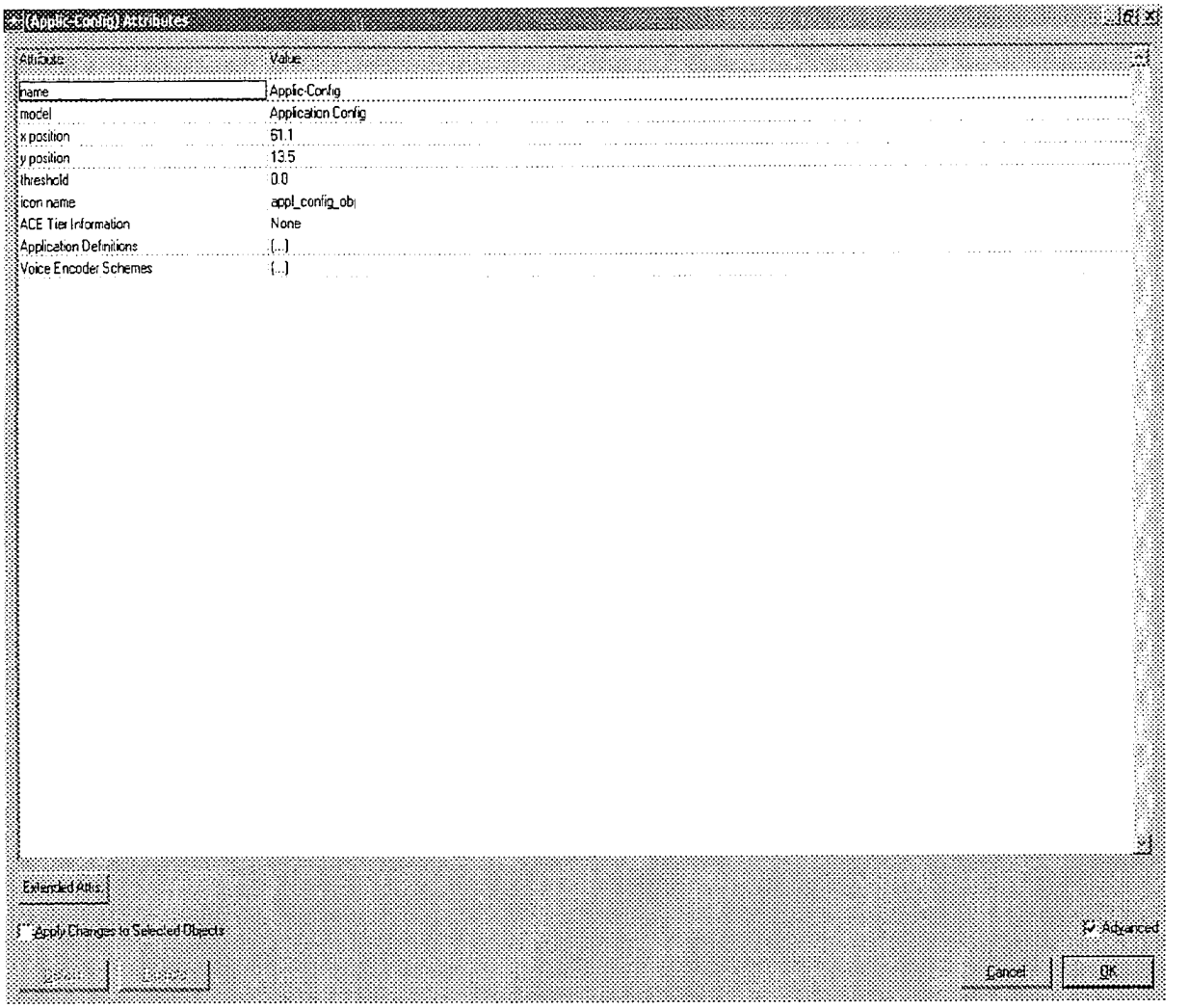


Figure 13.1. Application Configuration

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

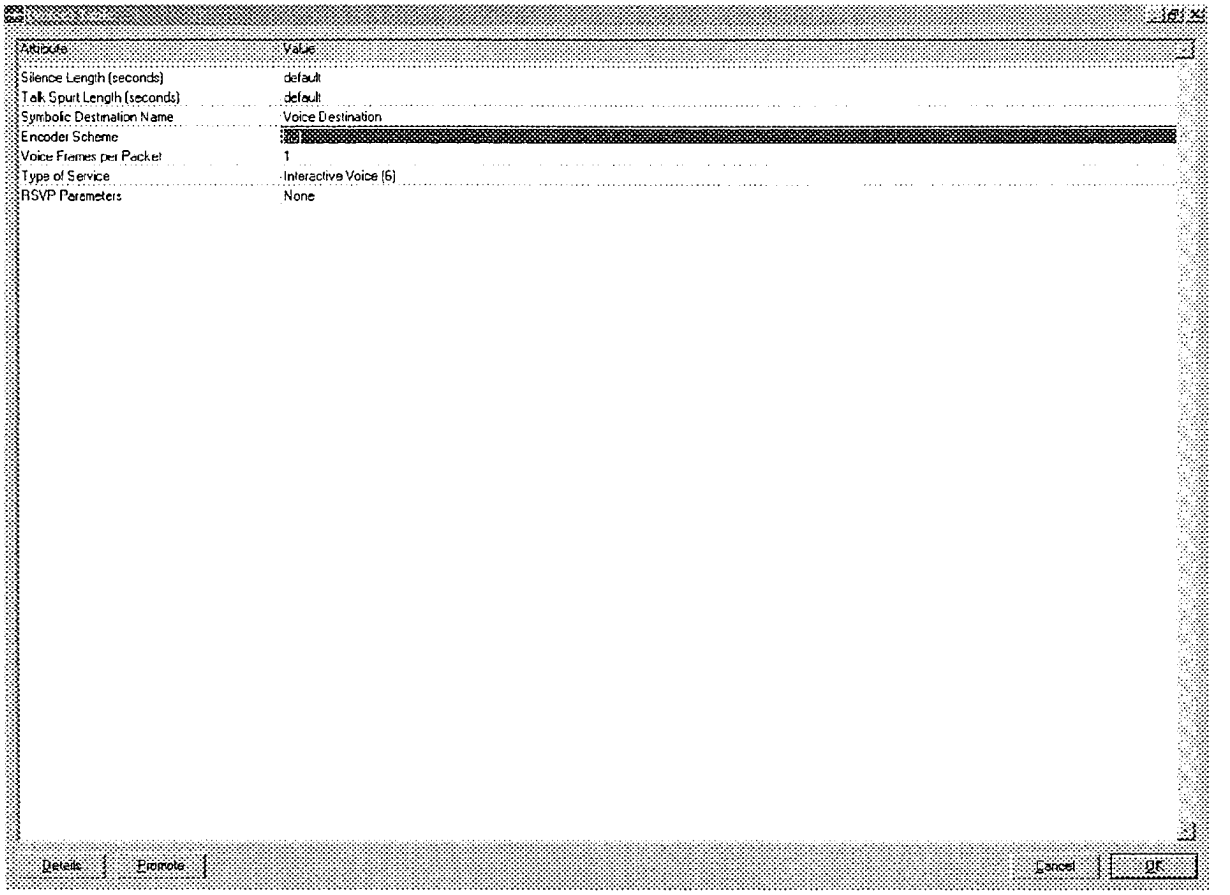


Figure 13.2. Application definition

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

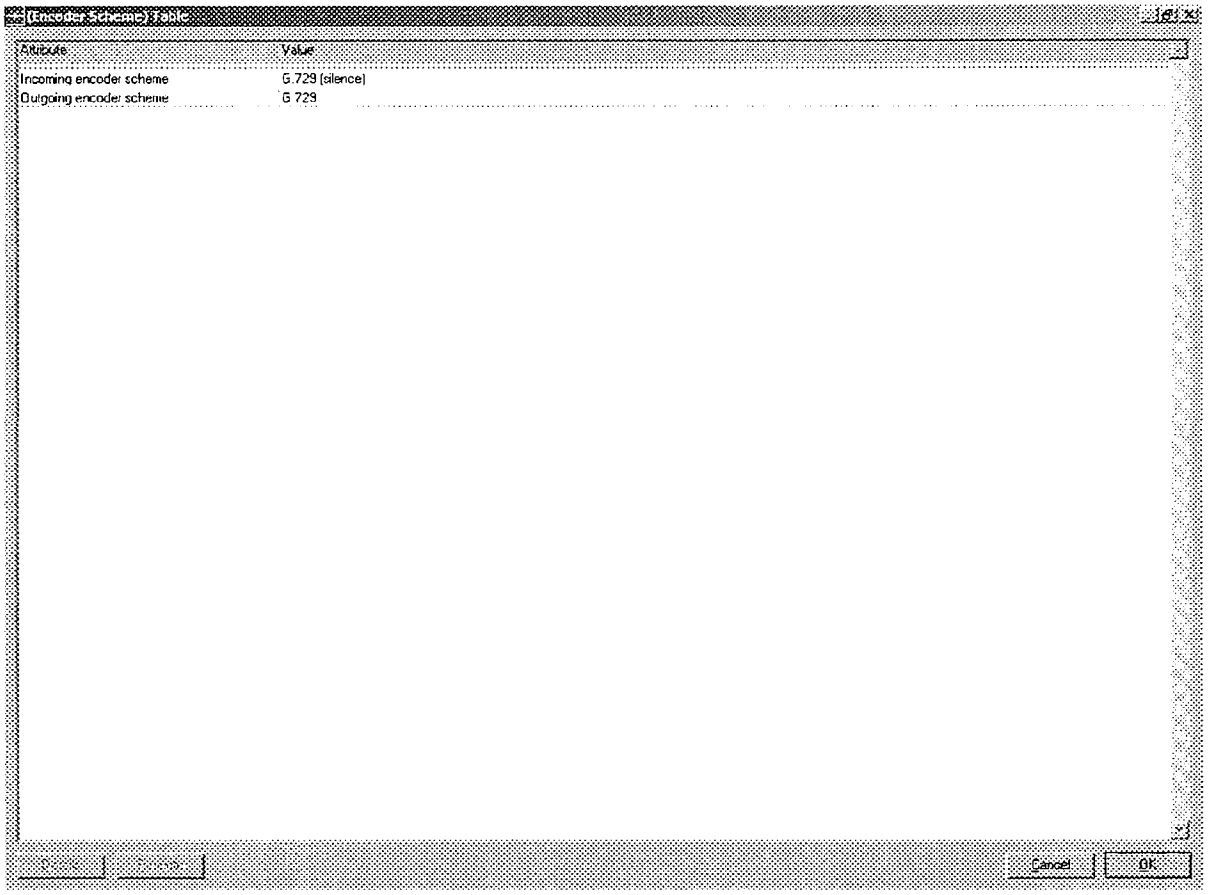


Figure 13.3. Encoder scheme

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

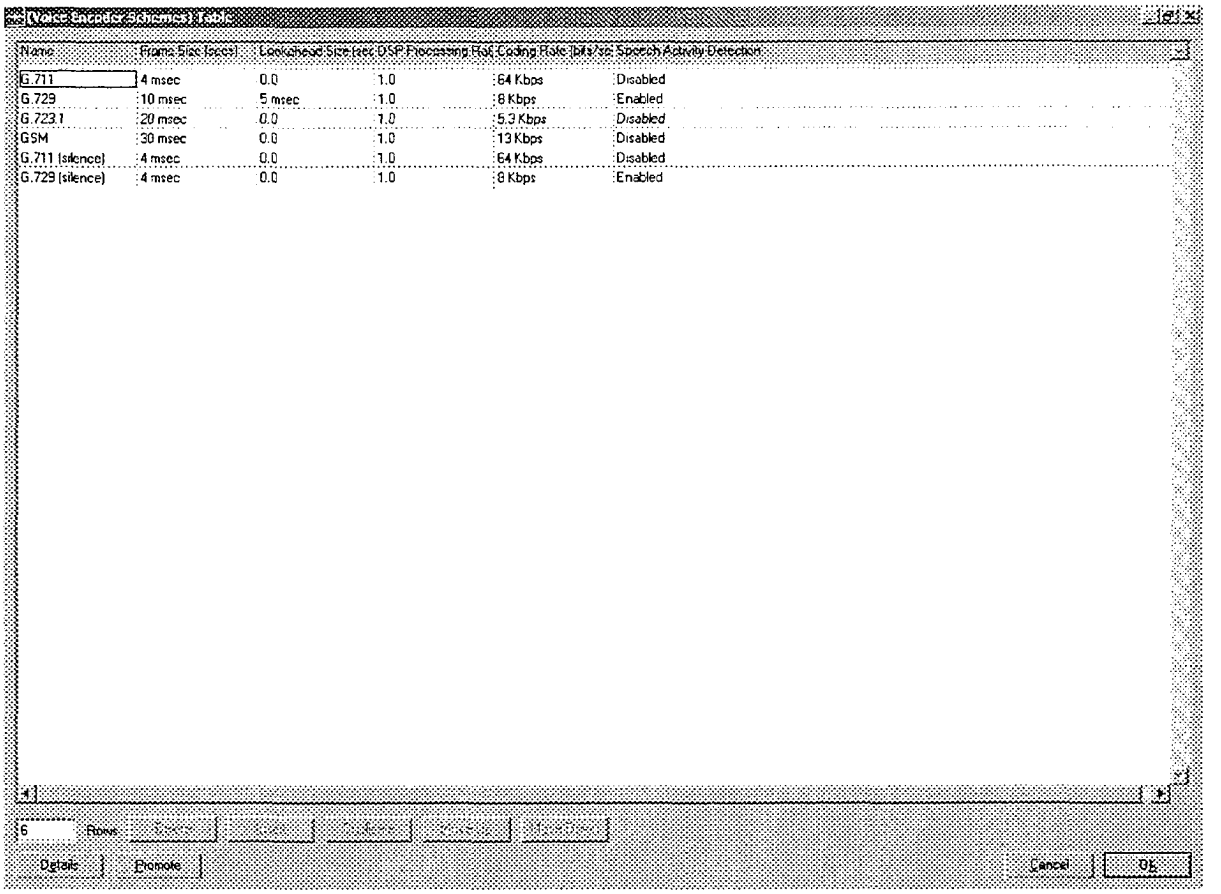


Figure 13.4. Implemented encoder schemes

II. Profile Configuration: specifies the symbolic name and the corresponding attributes for the defined profile. In order to configure a certain workstation to use a defined profile, it is necessary to set in the workstation, the attribute "Application Supported Profiles", to the defined value. We gave to the profile, the symbolic name of "Remote workstation". The Profile configuration has the following attributes:

1. Applications: specifies the different applications used in this profile and their attributes. These applications are defined in detail in the Application Configuration object.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

We have set this attribute to the value "Voice-IP Telephony", as defined in the applications configuration.

2. Operation Mode: specifies how the applications could start. Even if we have only one application the value is set to "Simultaneous", which gives the possibility that in case of several defined applications, all could start at the same time.

3. Start time: specifies the distribution name and parameters to be used for generating random outcomes for the profile start time. We set this attribute to a constant of mean outcome "0".

4. Duration: specifies the distribution name and parameters to be used for generating random outcomes for the profile duration. In our simulation this attribute is set to "End of Simulation".

5. Repeatability: specifies the parameters used to repeat the execution of this profile. In our simulations the attribute is set to the value "Once at start time".

For details related to the Profile configuration setting, see the figures below:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

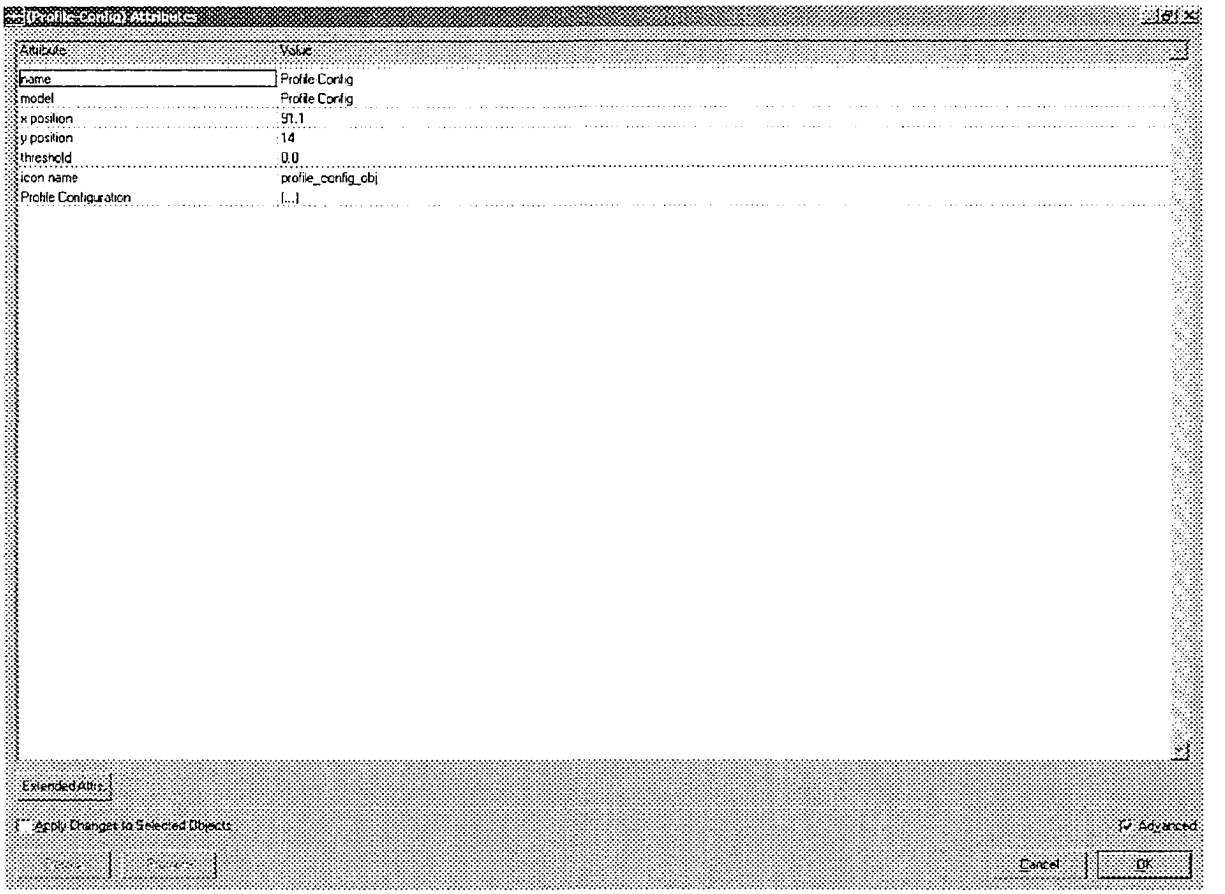


Figure 13.2.1. Profile Configuration

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

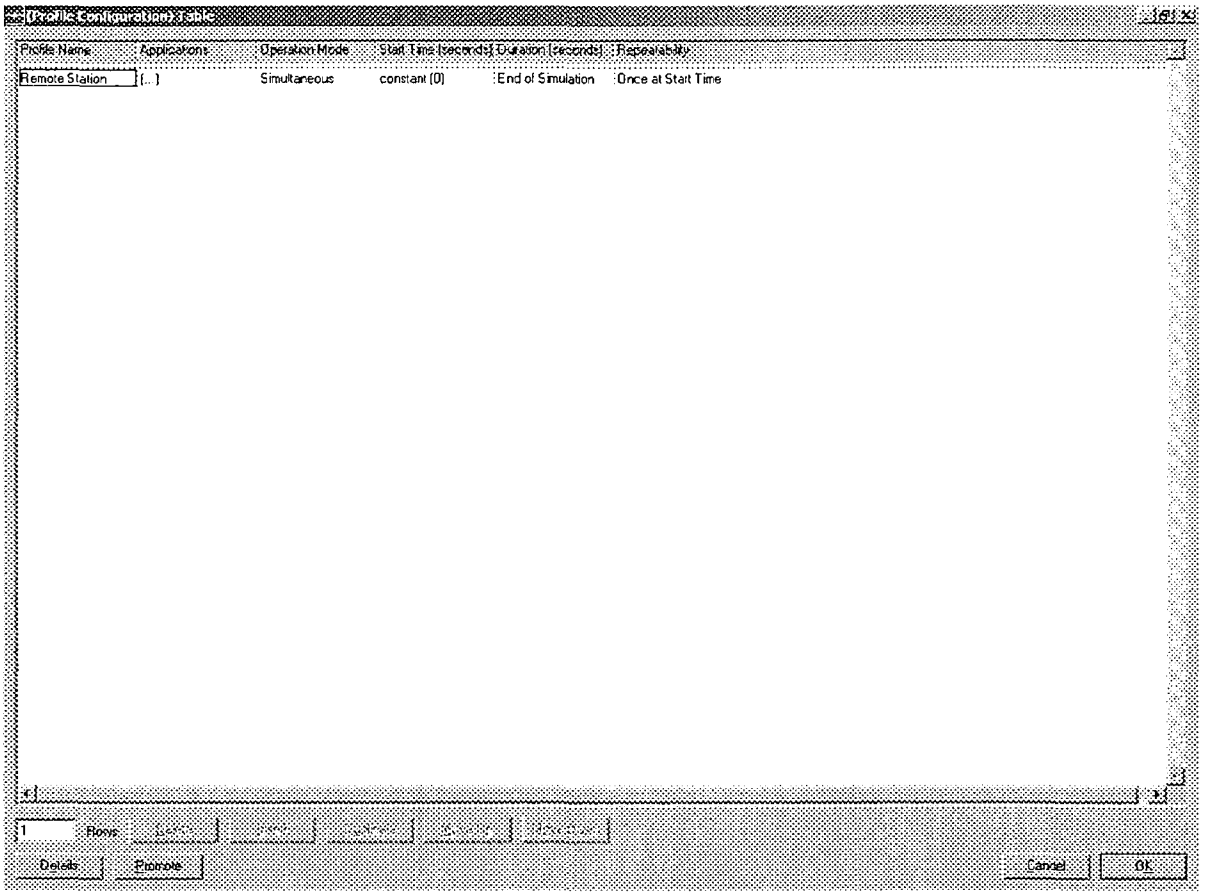


Figure 13.2.2. Profile configuration table

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

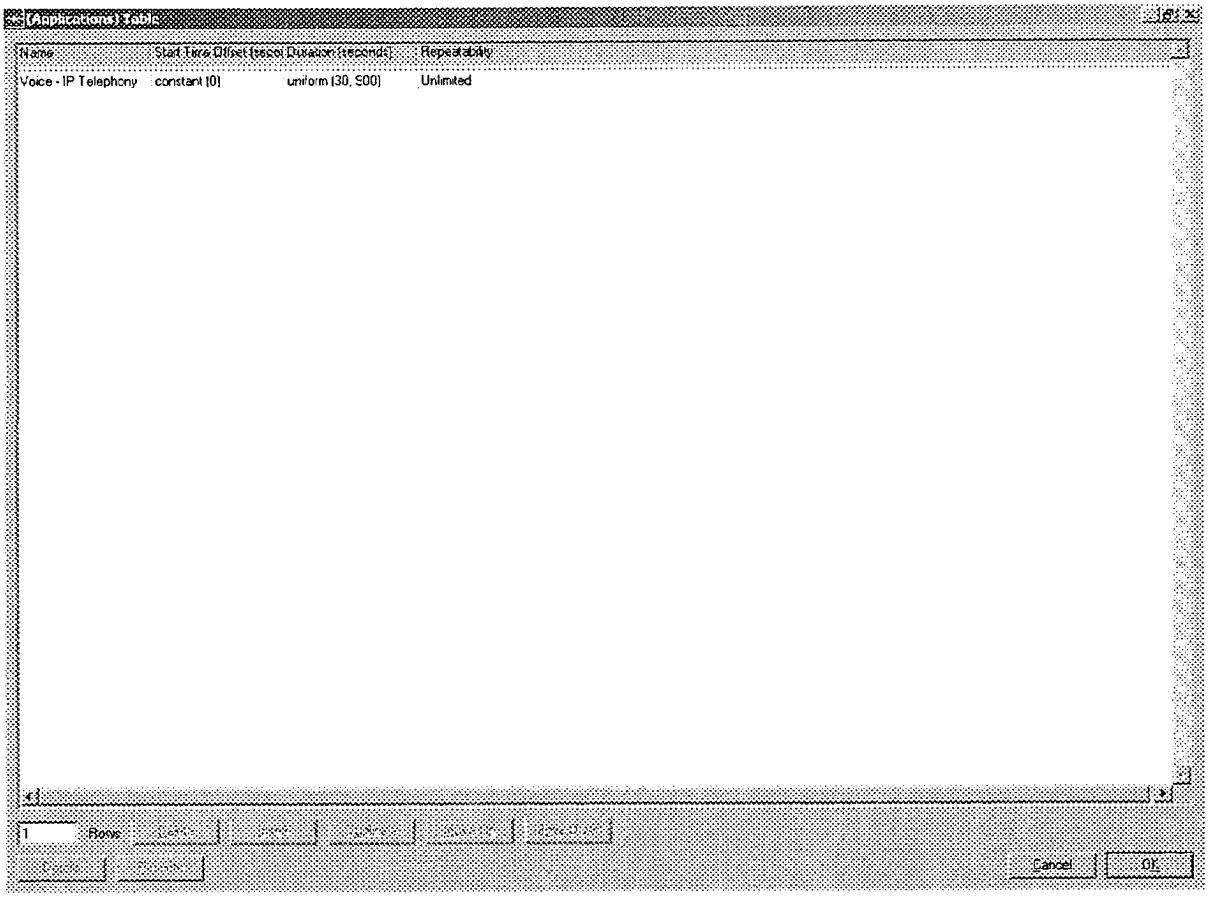


Figure 13.2.3. Implemented applications

III. IP Traffic Configuration: Defines attribute configuration details for the IP layer. These specifications can be referenced by the individual nodes using symbolic names . The following attributes are defined for IP traffic configuration:

1. IP Compression Information: Provides details of various compression schemes used in the network. In our simulation we use the Standard Compression scheme

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. IP Ping Parameters: specifies different ping option settings that individual hosts/routers in the network can use to determine connectivity to the specified destination. The following ping parameters are simulated:

- Interval: specifies the time in seconds between two consecutive ping packets; it is set to 1 sec.

- Packet size: specifies the size in bytes of the ping packet; it is set to 64 bytes.

- Count specifies the number of ping packets to be sent; it is set to 5.

-timeout: specifies the time in seconds, after which a ping packet is considered to be lost; it is set to 20 seconds.

- Record route: specifies the option of recording the route taken by the ping packet; it is set to "disabled".

For details related to the IP traffic configuration setting, see the figures below:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

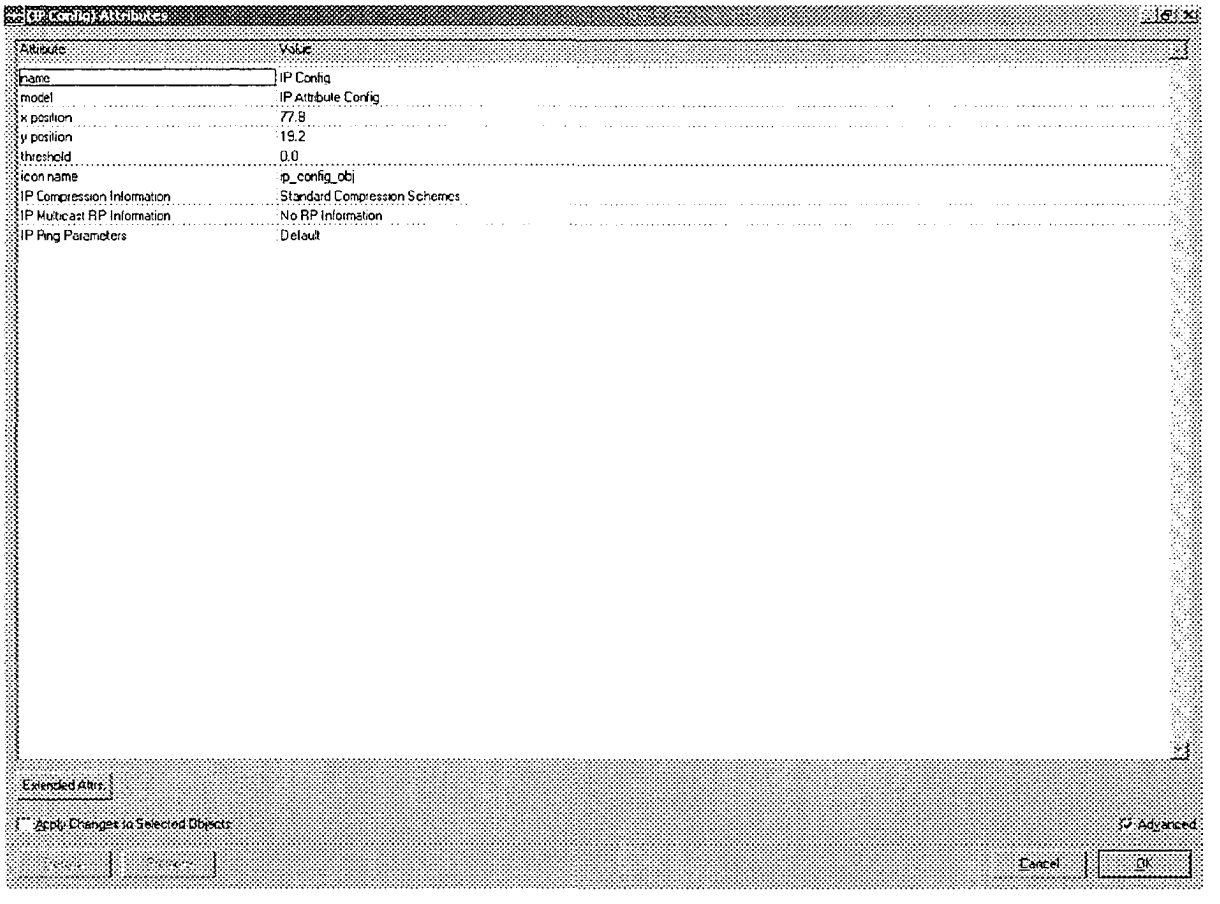


Figure 13.3.1 IP Traffic Configuration

IV. Quality of Service Configuration: defines the Quality of Service implemented through priority queuing or custom queuing.

Priority queuing could be based on

- Type of Service
- Protocol
- Port

We have implemented 8 levels of Type of Service and will compare how priority works between the lowest level, "Best Effort" and the highest level, "Interactive Voice".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As protocol based priority queuing we will use the RIP routing protocol.

We did not implement port based priority queuing.

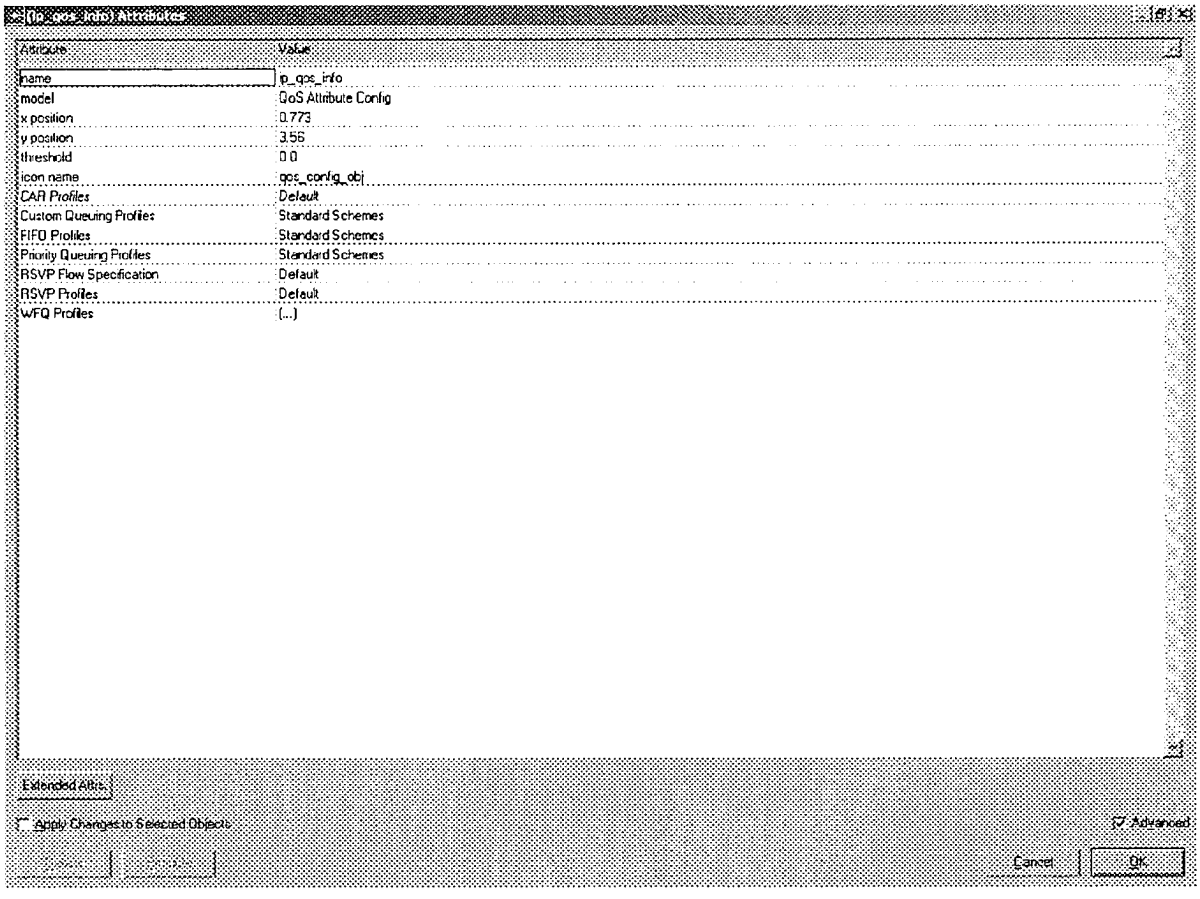


Figure 13.4.1 QoS Configuration

IV. Wireless workstation attributes: The wlan_wkstn_adv node model represents a workstation with client-server applications running over TCP/IP or UDP/IP.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The workstation supports one underlying WLAN connection at a transmission/reception speed of 1 Mbps, 2 Mbps, 5.5 Mbps, and 11 Mbps. We use the workstation simulation at a speed of 11Mb/sec.

This workstation requires a fixed amount of time to route each packet, as determined by the "IP Forwarding Rate" attribute of the node. Packets are routed on a first-come-first-serve basis and may encounter queuing at the lower protocol layers, depending on the transmission rates of the corresponding output interfaces.

The following protocols are implemented in the Fixed Wireless Workstation: RIP, UDP, IP, TCP, IEEE 802.11, OSPF

Implemented transmission speed are: 1 Mbps, 2 Mbps, 5.5 Mbps, 11 Mbps

The Fixed Wireless workstation has the following attributes:

1. Client Custom Application, Client Database Application, Client Email, Client Ftp, Client Remote Login, Client X Windows, Client Video Conferencing,

2. Client Start Time: These attributes allow for the specification of application traffic generation in the node.

3. Transport Address: This attribute allows for the specification of the address of the node.

4. IP Forwarding Rate: specifies the rate (in packets/second) at which the node can perform a routing decision for an arriving packet and transfer it to the appropriate output interface.

5. IP Gateway Function: specifies whether the local IP node is acting as a gateway. Workstations should not act as gateways, as they only have one network interface.

6. RIP Process Mode: specifies whether the RIP process is silent or active. Silent

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

RIP processes do not send any routing updates but simply receive updates. All RIP processes in a workstation should be silent RIP processes.

7. TCP Connection Information: specifies whether diagnostic information about TCP connections from this node will be displayed at the end of the simulation.

8. TCP Maximum Segment Size: determines the size of segments sent by TCP. This value should be set to largest segment size that the underlying network can carry unfragmented.

9. TCP Receive Buffer Capacity: specifies the size of the buffer used to hold received data before it is forwarded to the application.

Below we present the traffic flow for the various layers of the implement model.

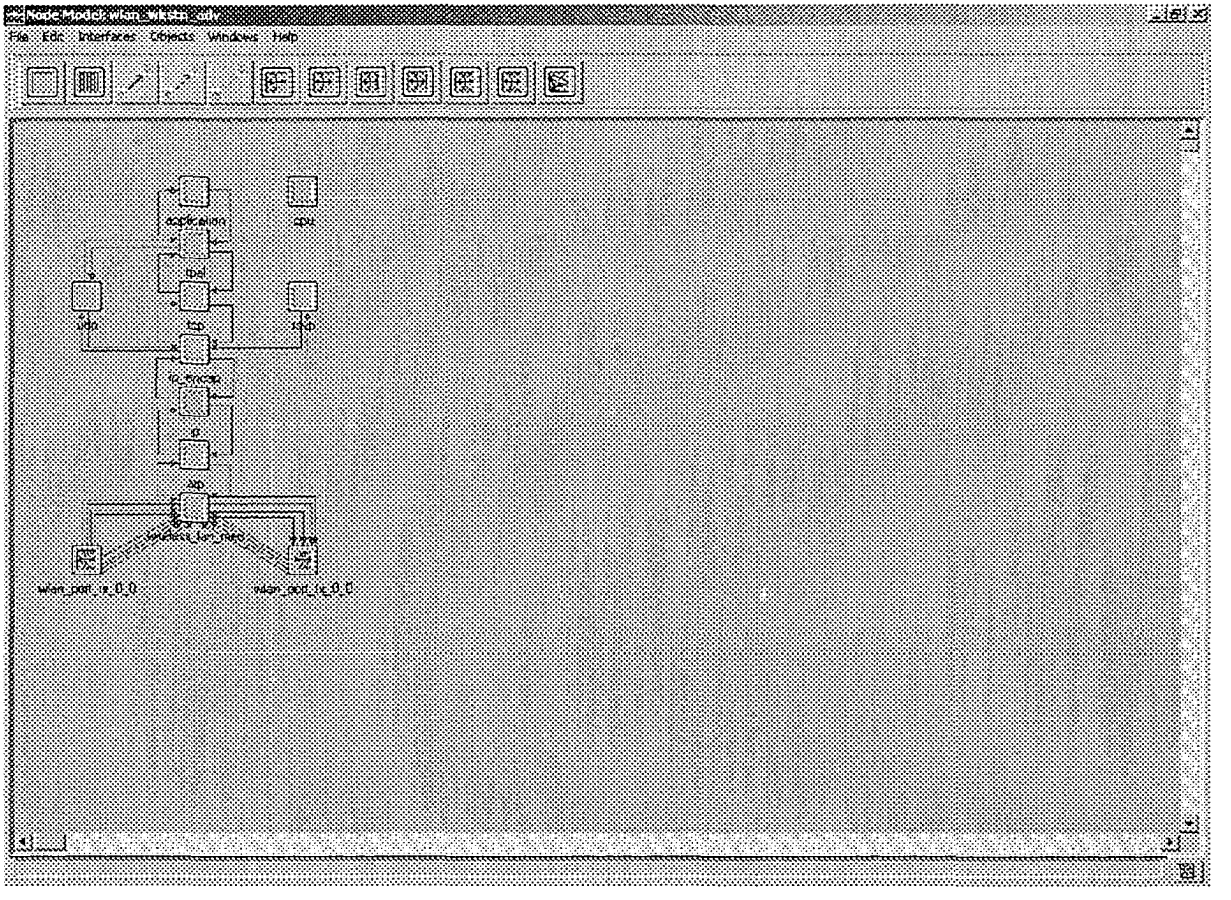


Figure 13.5.1 General Node Model

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. Application layer:

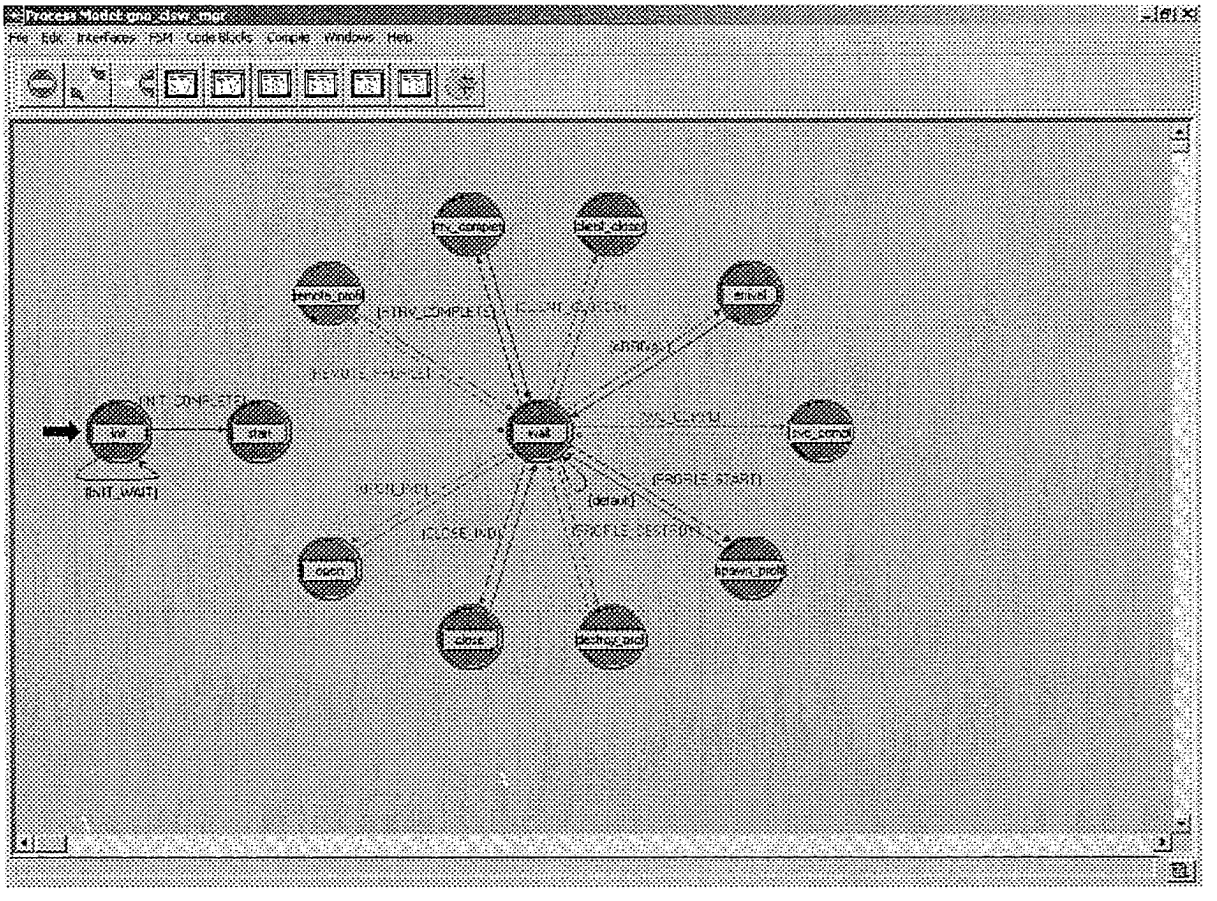


Figure 13.5.2 Traffic Flow at Application Layer level

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. TPAL layer:

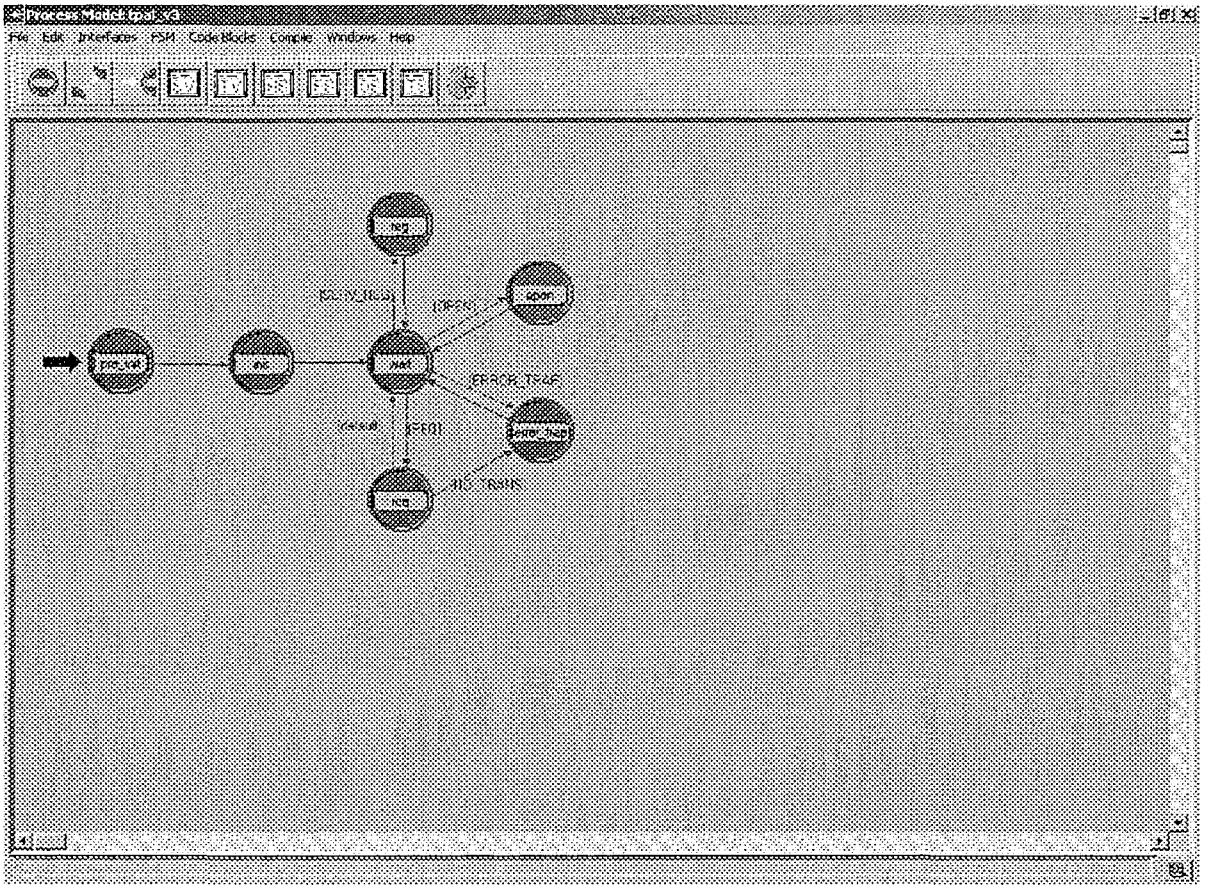


Figure 13.5.3 Traffic flow at TPAL Level

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4. TCP layer:

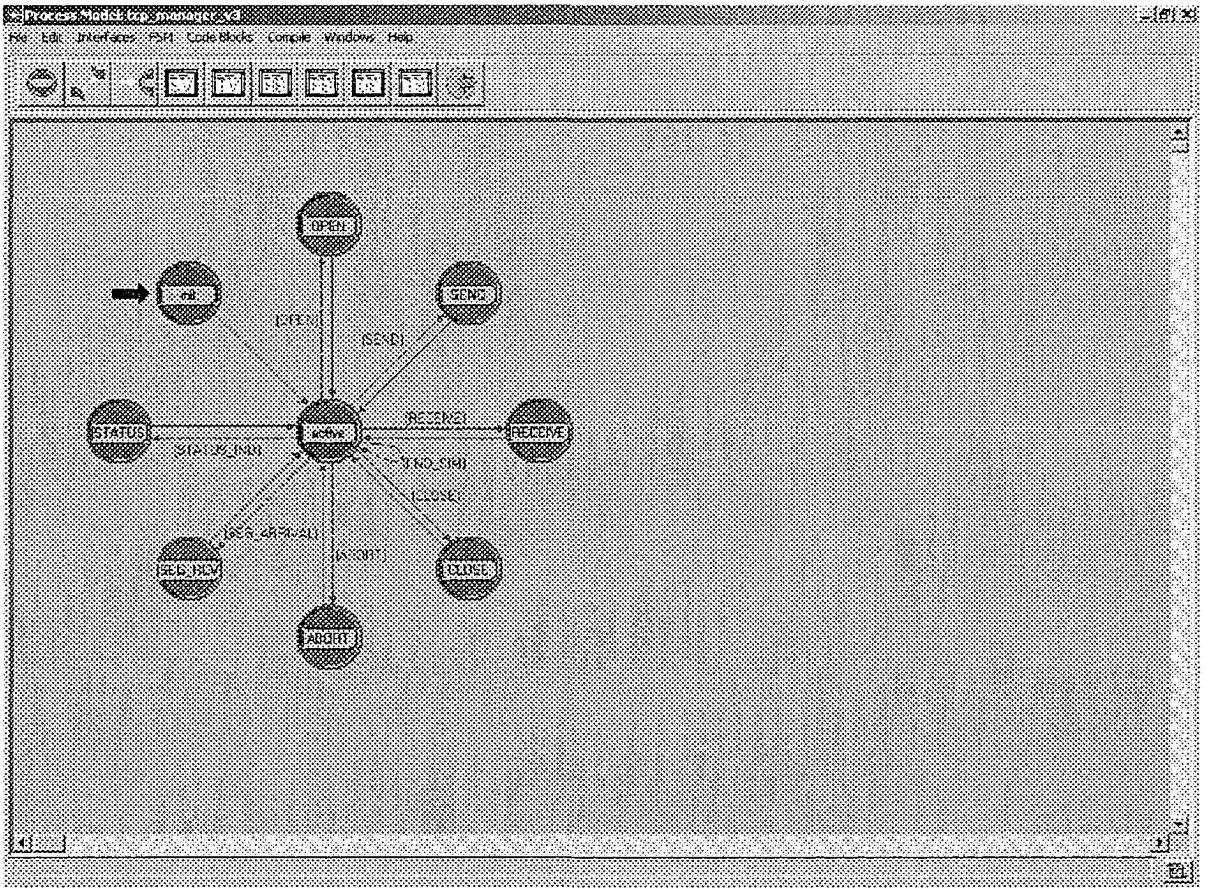


Figure 13.5.4 Traffic flow at TCP Layer Level

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4' UDP layer:

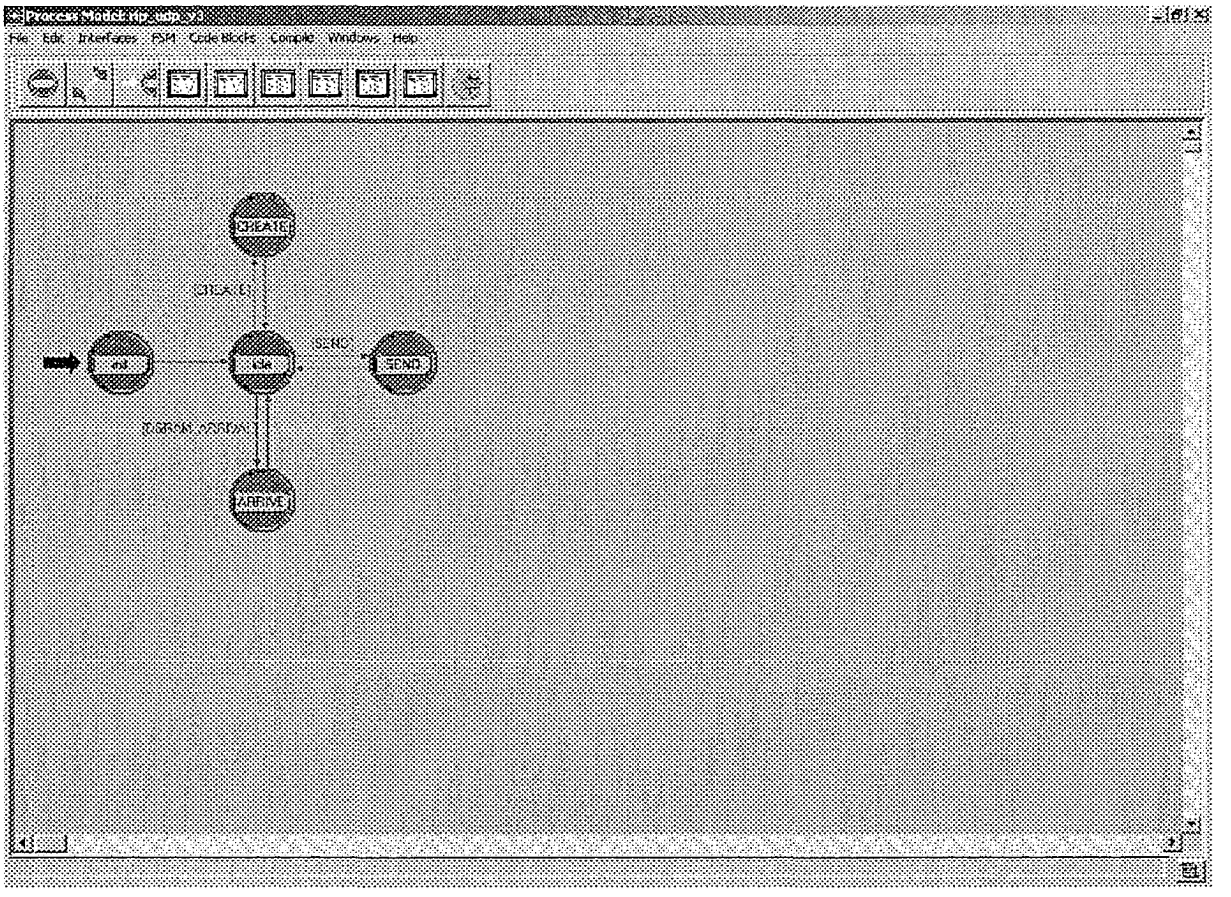


Figure 13.5.4' Traffic flow at UDP Layer Level

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5. IP encapsulation:

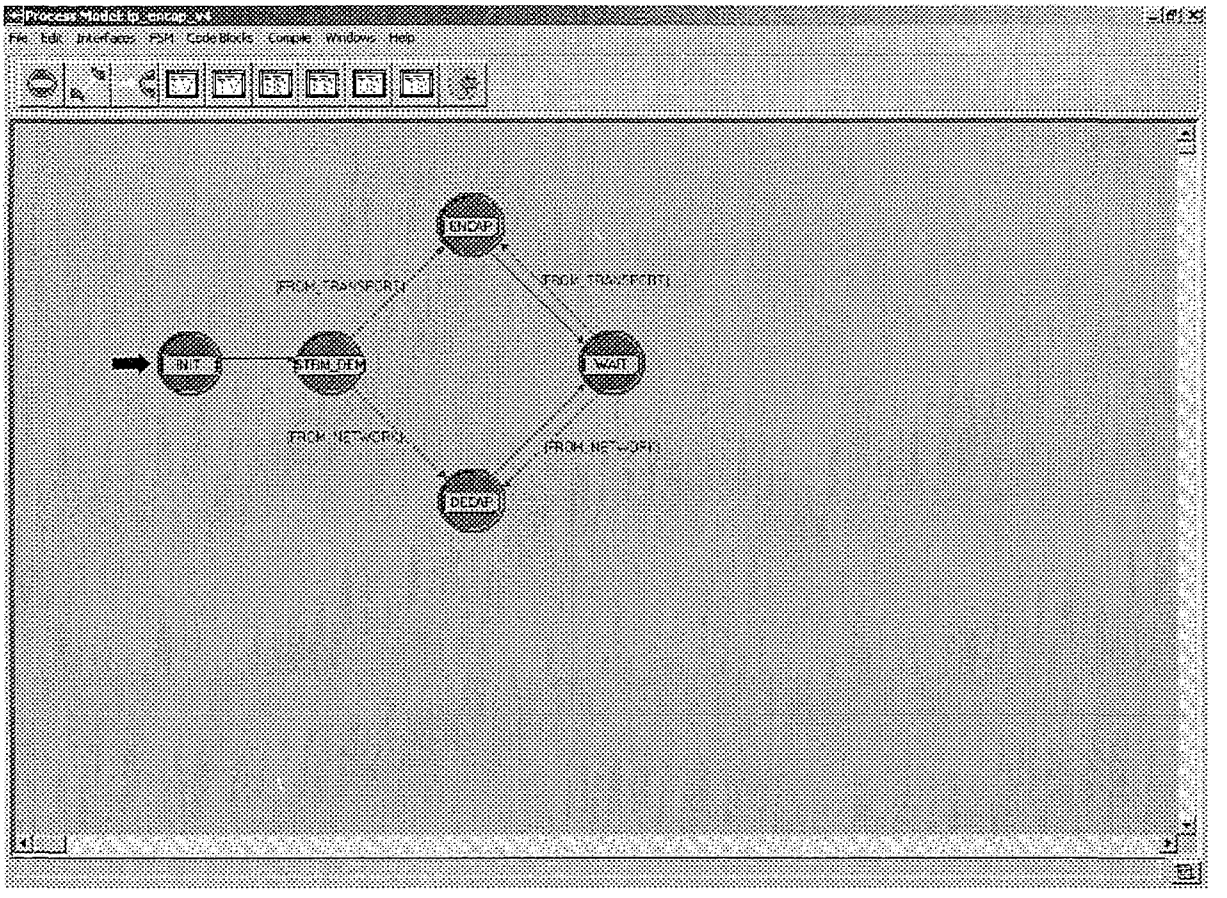


Figure 13.5.4 Traffic Flow at IP Encapsulation Level

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5'. IP layer:

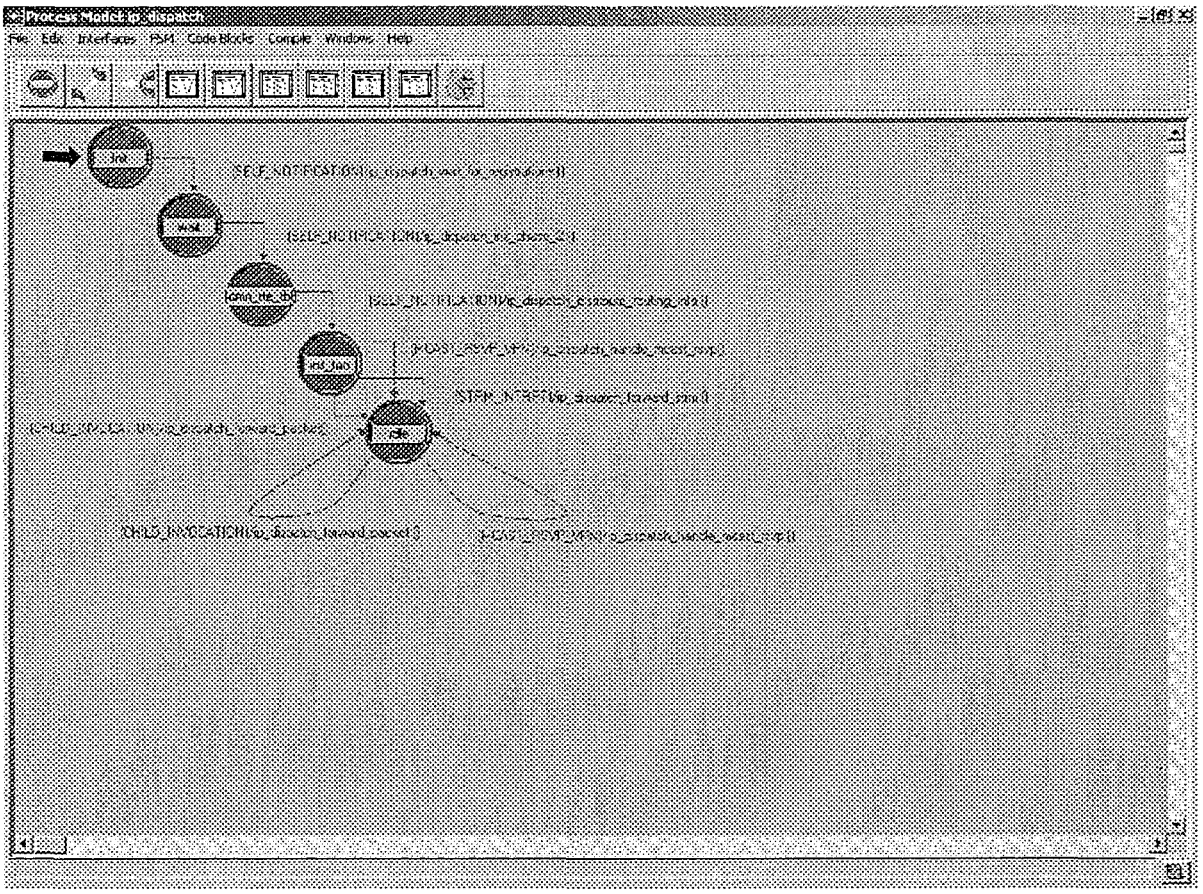


Figure 13.5.5' Traffic Flow at IP Layer Level

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5. ARP level:

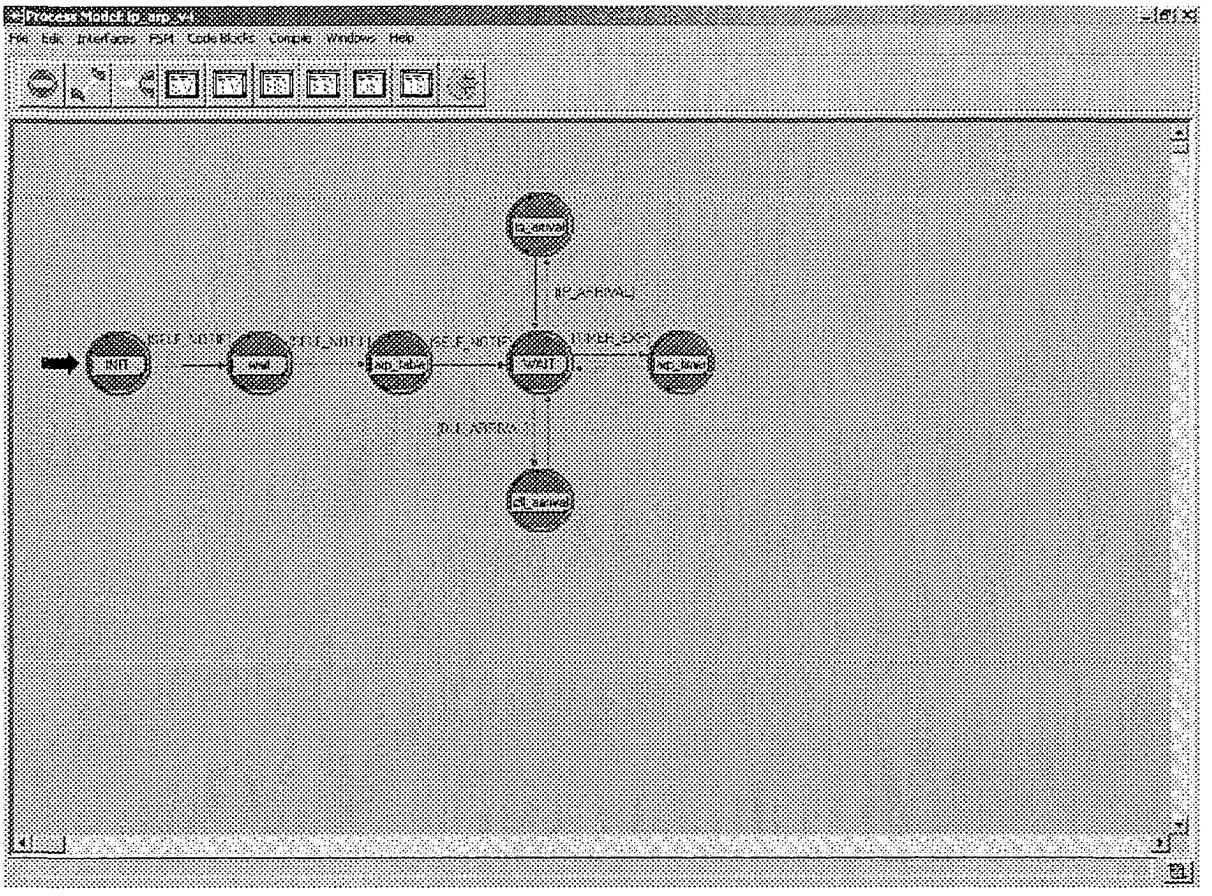


Figure 13.5.5' Traffic Flow at ARP Level

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6. WLAN data link layer:

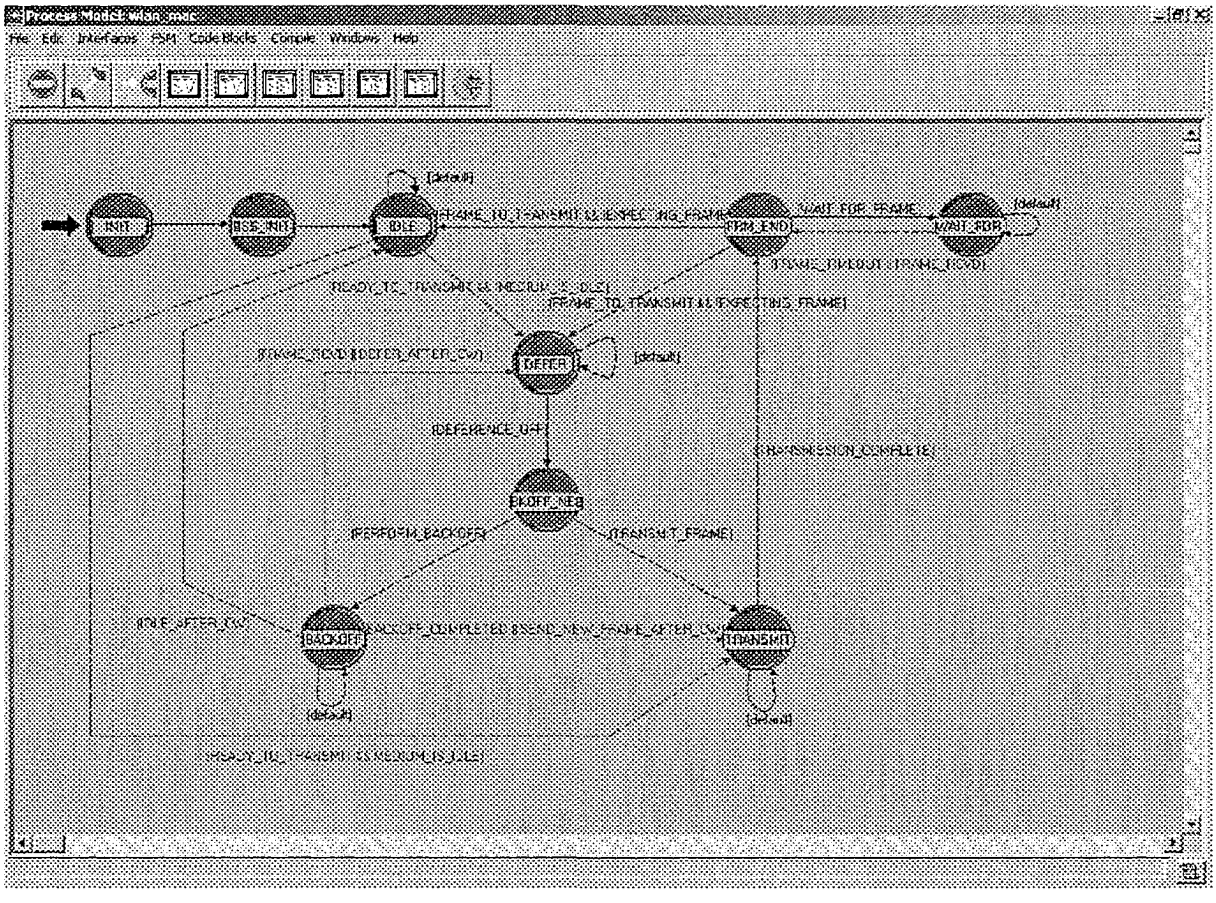


Figure 13.5.5' Traffic Flow at Data Link Layer Level

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Node Model contains the following attributes, with corresponding setting:

- Application destination preference: provides the mapping between symbolic destination names specified in the Application Definition object and the real name specified in the Server name or Client Name of each node. Each symbolic destination can map a set of real destinations, and this destination is chosen based on a weight value. The weight in fact specifies the probability with which the destination is chosen.
In our case The "Voice destination" is mapped to a wireless workstation . The chosen numbering scheme is in fact the numerical order of the workstation, in a sequential way.

- Segment size: is the maximum size in bits of a segment passed on to the TPAL layer, from the application. This attribute is set to the value of 63Kb and will not be changed during simulations.

- Application supported profiles: specifies the name of all profiles enabled on the specified node. The profile is fully defined in the Profile Configuration object. A profile describes the user behavior in terms of what applications are used, the traffic of the application, duration, use and repeatability. The name of used profile in the simulations is "Remote Station" as defined in the Profile Configuration.

- Application Transport Protocol Specification: specifies the transport protocol being used by each defined application. For Voice application the UDP protocol is set.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- CPU background utilization: specifies the average baseline loads for the given link. For each packet that is transmitted through the link a delay is calculated. This is in addition to the background traffic specified in the Traffic Browser.

- CPU Resource Parameters: defines the number of processors in the workstation. We use only 1 CPU workstations.

- Client address: specifies the TPAL layer level addressing. It is a unique number attached to each workstation. In our simulations represent the phone number of the workstation.

- TCP Parameters: specifies the parameters of the simulated TCP protocol. In our simulation the Voice application is not using TCP as a transport protocol, UDP being used, instead.

- Wireless MAC address: specifies the Medium Access Control layer address. It is a unique address and in our simulations is "Auto Assigned" by the simulator.

- Wireless LAN Parameters: specifies the parameters used by the Mac layer, WLAN protocol. It has set the following sub-attributes:

  - RTS threshold is set to "None"
  - Fragmentation threshold is set to "None
  - Data rate is set to 11Mbps
  - Physical characteristics: is set either to "Direct Sequence or to Frequency Hopping

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Short retry limits is set to 7

- Long retry limits are set to 4

- Access point functionality is disabled

- Channel settings are:

- Bandwith of 10KHz, and

- Minimum Frequency is set as a function of the BSS identifier in the node, using the formula:

Freq= (BSS_identifier +2)*30, so in our case BSS_identifier=1 so

Freq=30MHz

- Buffer size is set to 256Kb

- Max. receive lifetime is set to 0.5 seconds

- Large packet processing is set ot "Drop packet"

- BSS identifier is not used in ad-hoc and infrastructure networks and is set to consecutive .

- Financial Cost specifies the cost of purchasing and maintaining the network; it is not used in our simulations.

The figures below show the settings for the parameters described above:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

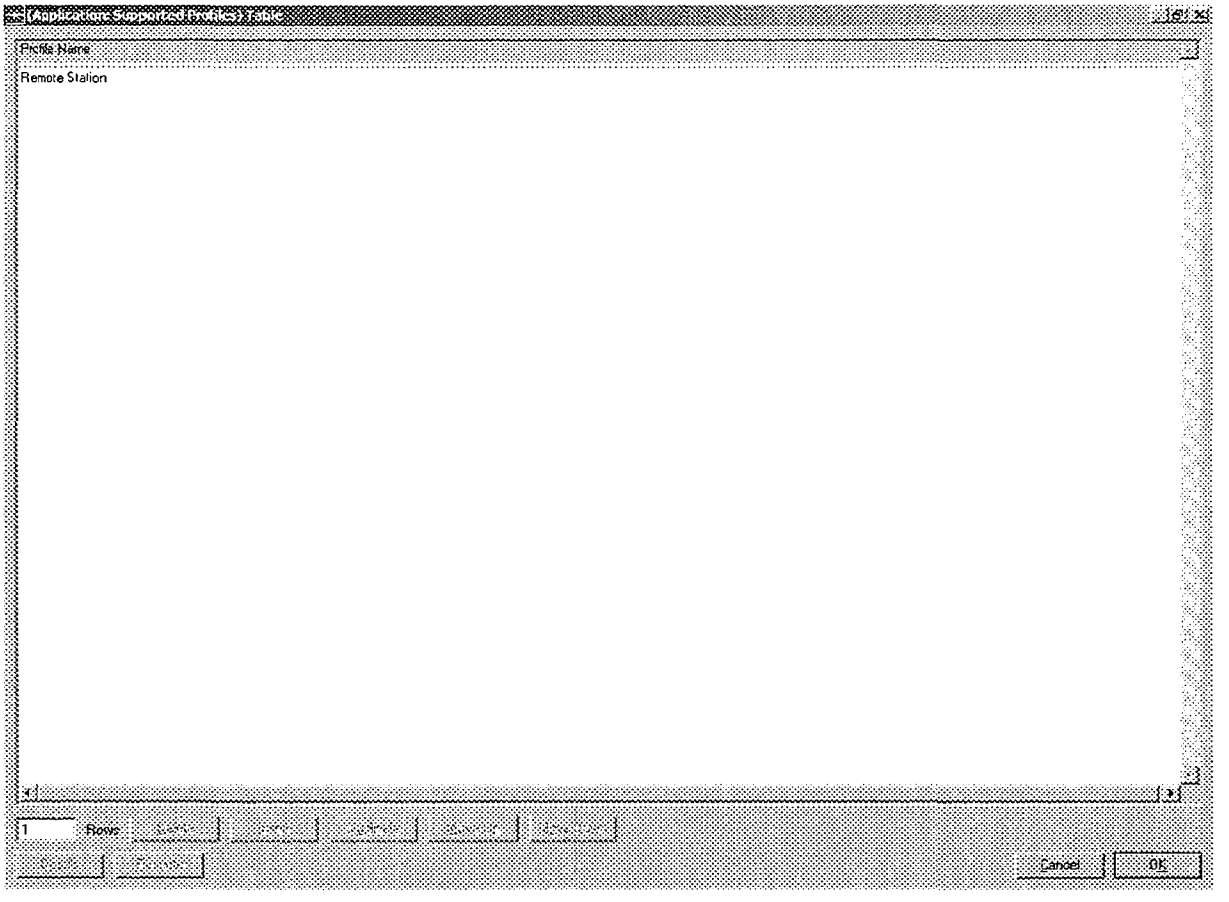


Figure 13.6.1 Supported profiles of the application:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

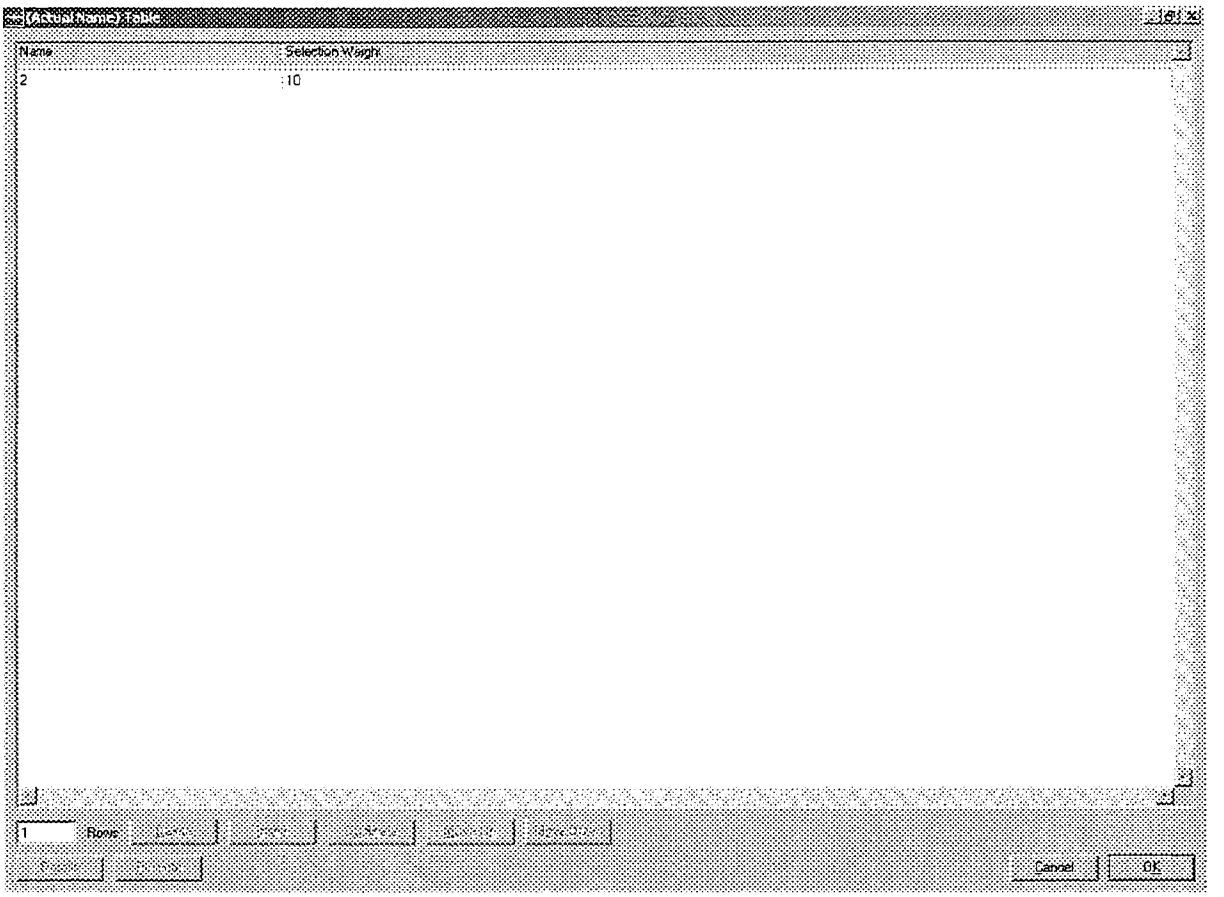


Figure 13.6.2 Calling Address

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

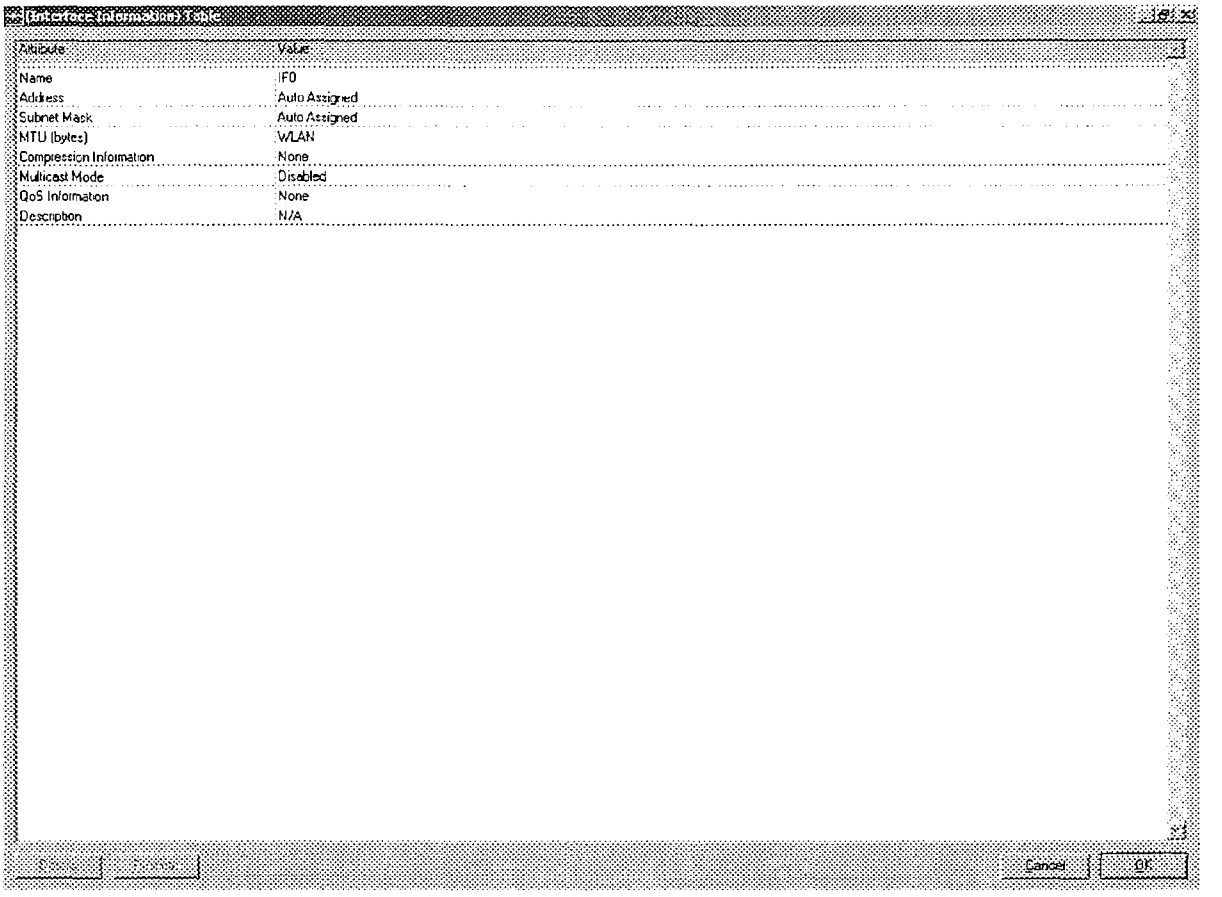


Figure 13.6.3 Interface information

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

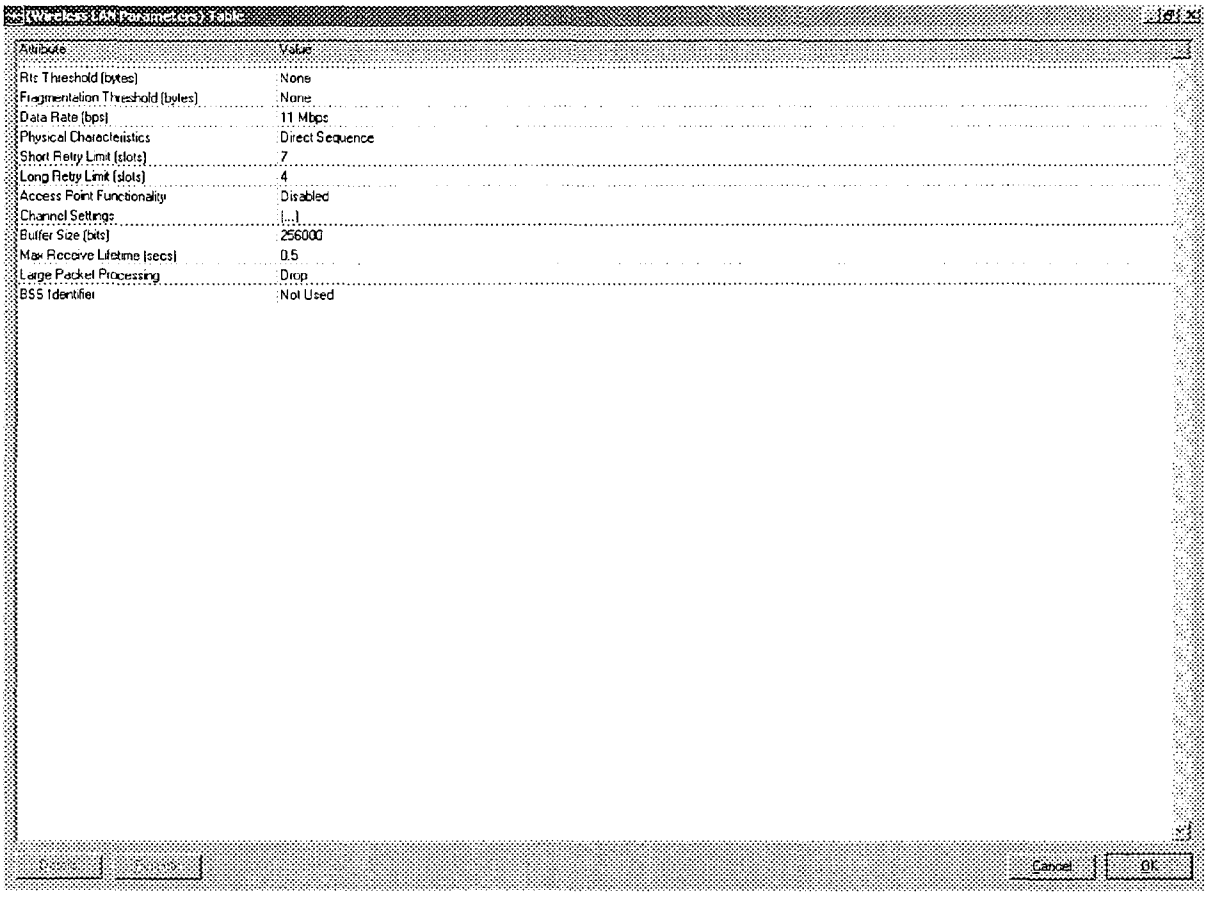


Figure 13.6.4 Wireless LAN parameters

The channel settings are:

- Bandwith: is set to 10KHz
- Frequency: is set to 30MHz)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

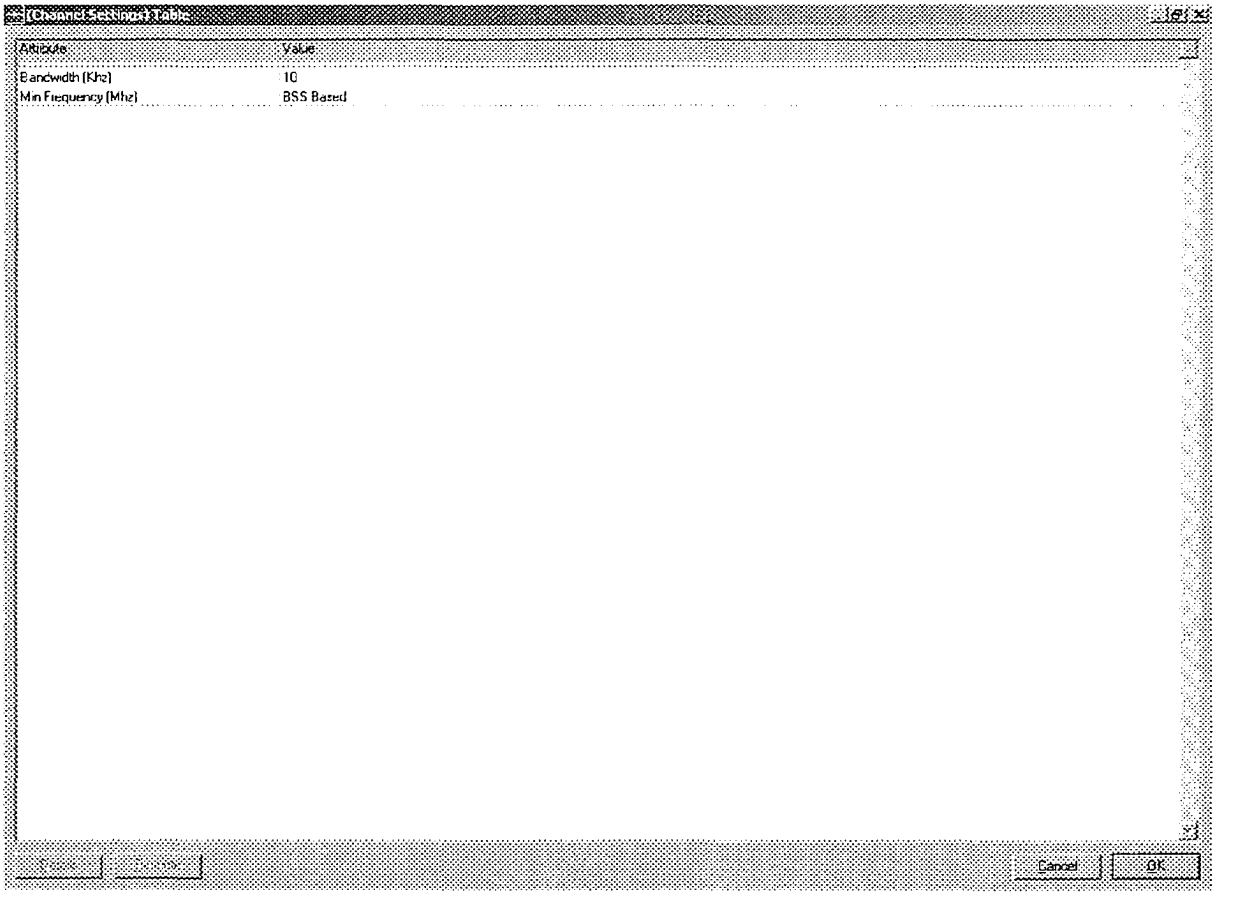


Figure 13.6.5 Channel setting

Simulation set-up an running

In order to run the simulation, the following settings are necessary:

Duration: in some cases, 30 minutes of network activity is simulated, but in other complex scenarios it was not possible to simulate more than 10 minutes of network activity. For each scenario the running time will be stated.

Note: In complex architectures, 10 minutes of network activity required 8-10 hours of running time and as much as over 750 billion events were generated !

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Values per statistic: represent the number of values collected per simulation for each statistic. For most of the scenarios we collect 100 values per statistic, but in case of complex scenarios we lower it to as small as 15. For each scenario the value is stated. Seed represents the random number generation seed and is varying from one simulation to the other.

Simulation attributes represent the simulation attributes relevant to each scenario.

The following attributes are set:

IP Interface Addressing Mode specifies the IP interface addressing mode. It is set to "Auto addressing", which means that all IP interfaces are assigned IP addresses automatically during simulation. The class of address (e.g., A, B, or C) is determined based on the number of hosts in the designed network. Subnet masks assigned to these interfaces are the default subnet masks for that class.

Wireless LAN range is set to 300 meters in each direction.

Generate web report is enabled, allowing the system at the end of simulation to generate a report with the finding in html format. The most important results are copied form this web report.

Define statistic report, let the user to define what parameters to have statistics collected and in what format. We defined a special Wireless Voice Report, which contains statistics collected for the following simulated parameters:

- Global statistics related to voice:
- Packet delay variation
- Packet end-to-end delay
- Traffic received for all workstations

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Traffic sent for all workstations
- Global statistics for the Wireless LAN:
- Data dropped
- Delay
- Load
- Throughput

In some cases we collected individual statistics for each workstation in order to study the performance at individual level. For more complex simulations this was not possible due to the limitation of the PC.

For each scenario the statistic parameters are fully described and the results analyzed.

The statistic report was associated in each case with the running simulation scenario.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

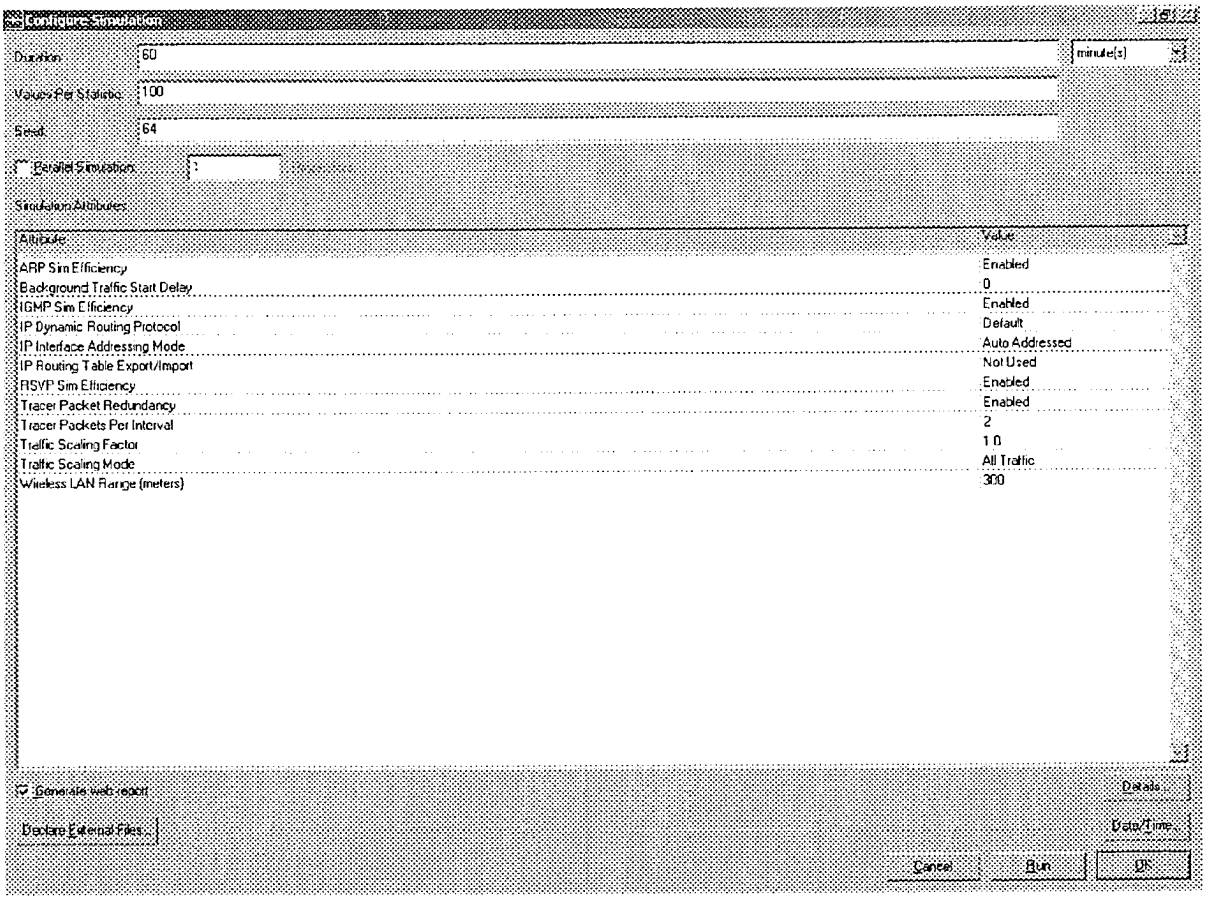


Figure 13.7.1 Simulation Configuration

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

13.4 Simulation Scenarios

I. Ad-hoc Networks

As already described in detail, ad-hoc networks do use peer-to-peer connection.

1st Scenario

In this scenario we study which physical layer protocol is more performant for voice transmission.

Having this goal in mind, we set up a Wireless LAN with the following characteristics:

- 802.11b at 11Mbps
- ad-hoc network without Access Point
- 32 fix wireless workstations
- voice codec G.729
- calling scheme is as follows: workstation-1 calls workstation-2, workstation 3 calls workstation-4, ..., workstation31 calls workstation32

We study which physical layer protocol is more performant for voice transmission over a wireless network: Direct Sequencing or Frequency Hopping.

The architecture of the simulation is presented in Figure1.1:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

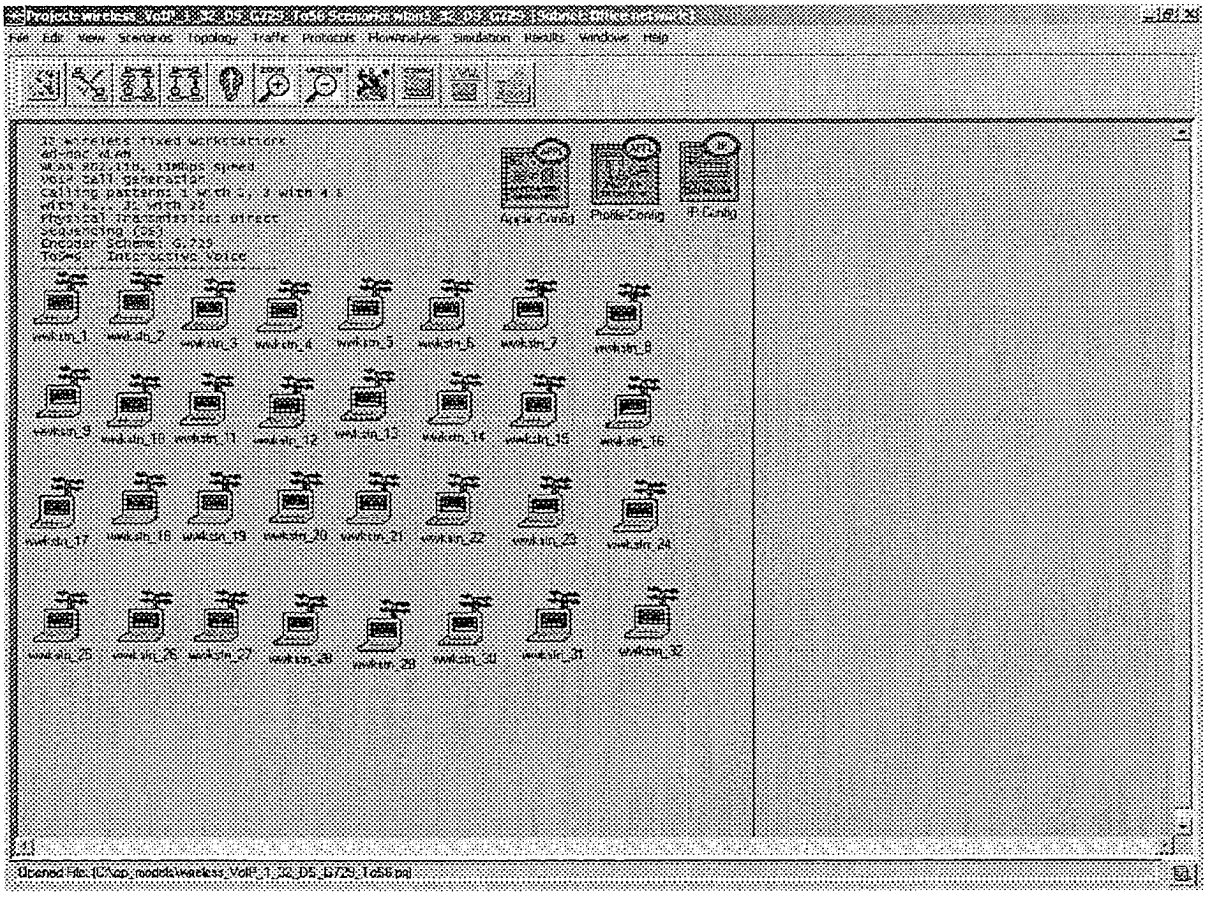


Figure 1. Architecture of ad-hoc wireless network with 32 users

We ran the simulation for 30 minutes.

The simulation results are:

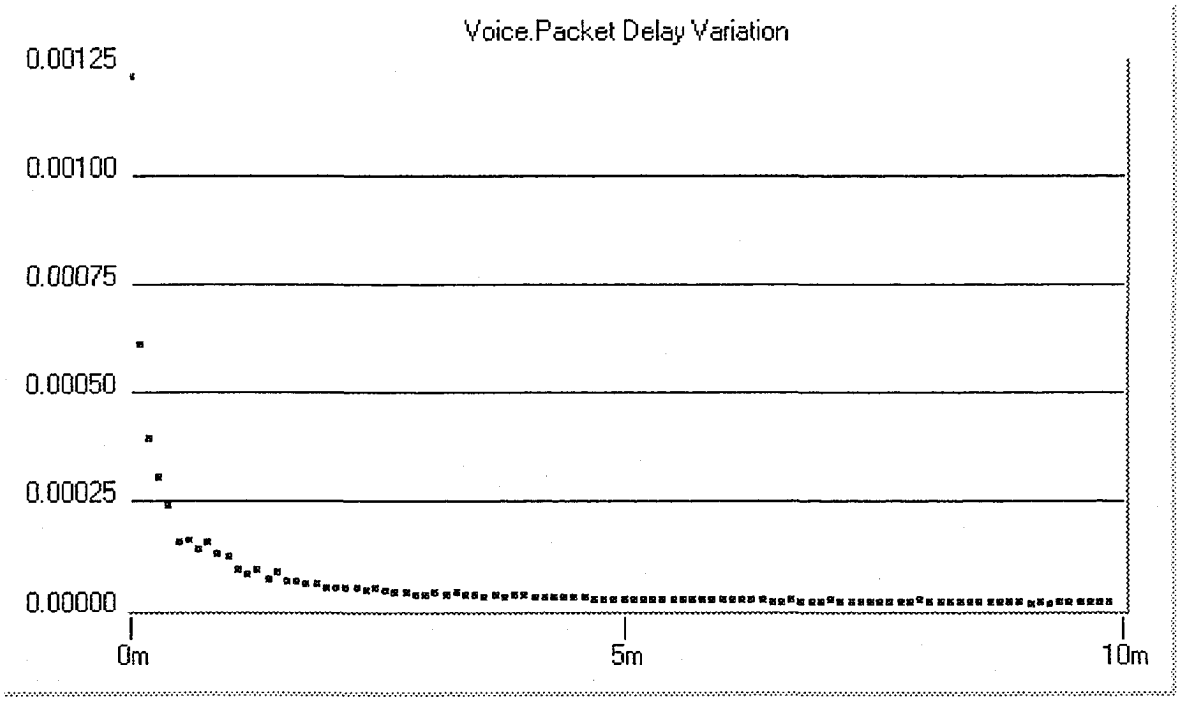
1.1.1 **For Direct Sequencing** the voice quality simulation results are:

- Voice quality parameters are:

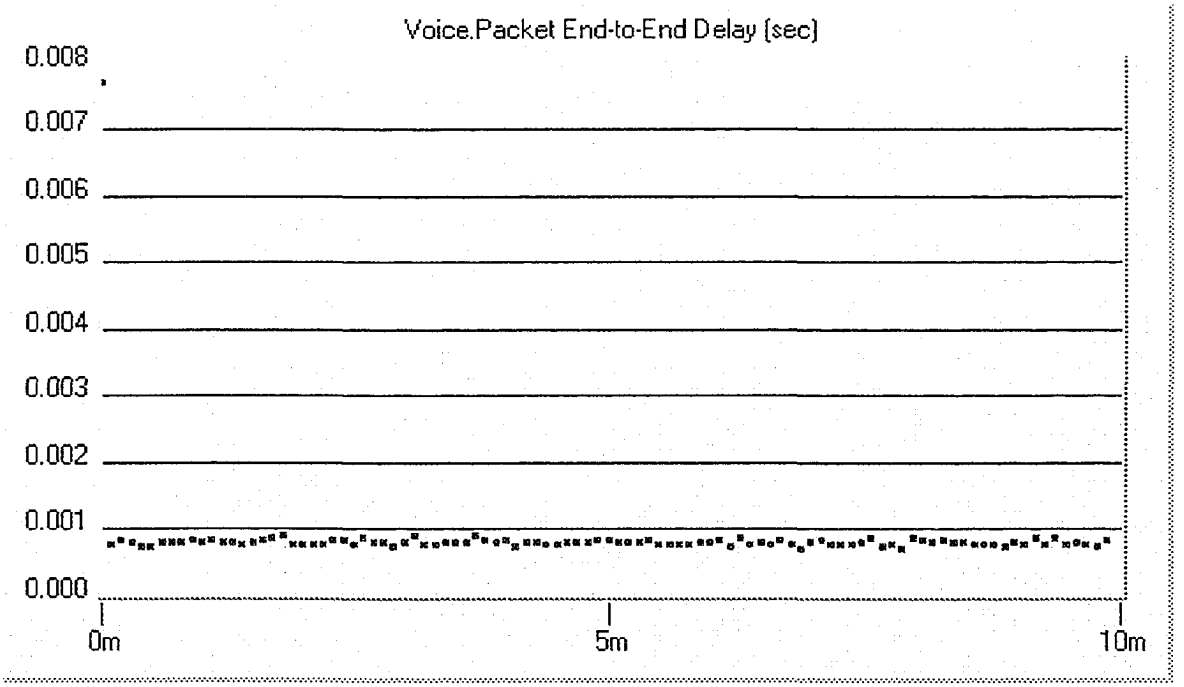
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.00006	0.00122	0.00001
Voice Packet End-to-End Delay (sec)	0.00083	0.00768	0.00067
Voice Traffic Received (bytes/sec)	17,596	19,480	15,592
Voice Traffic Sent (bytes/sec)	17,599	19,741	15,593

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graphical representation for voice packet delay variation:

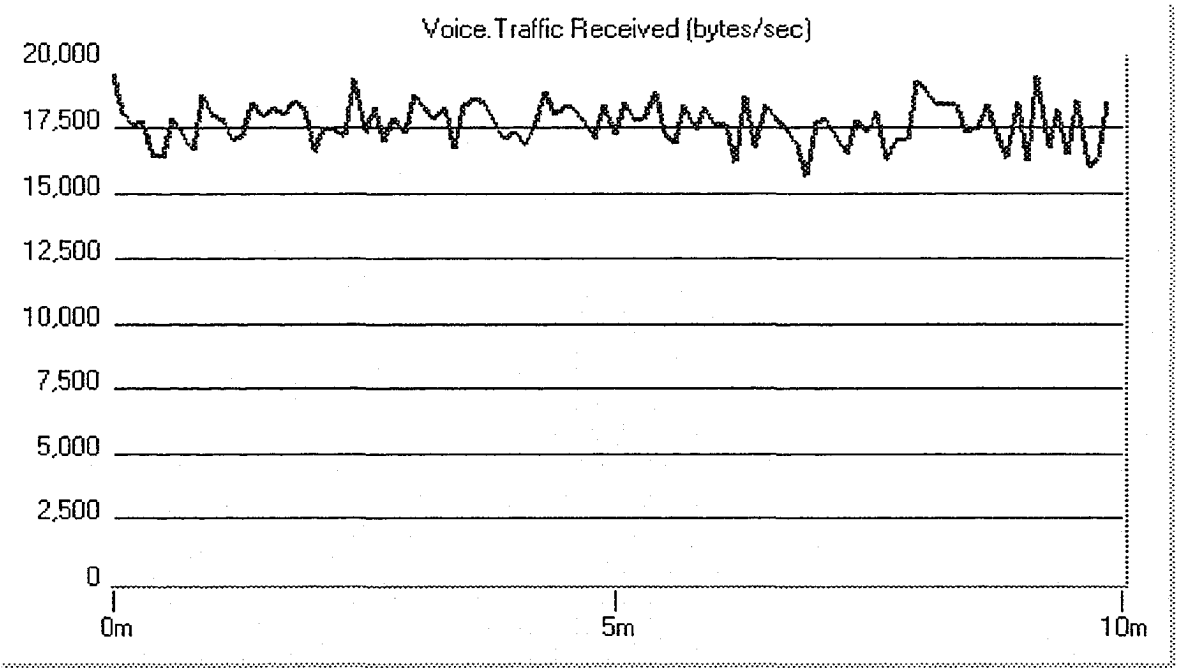


- Graphical representation for voice packet delay:

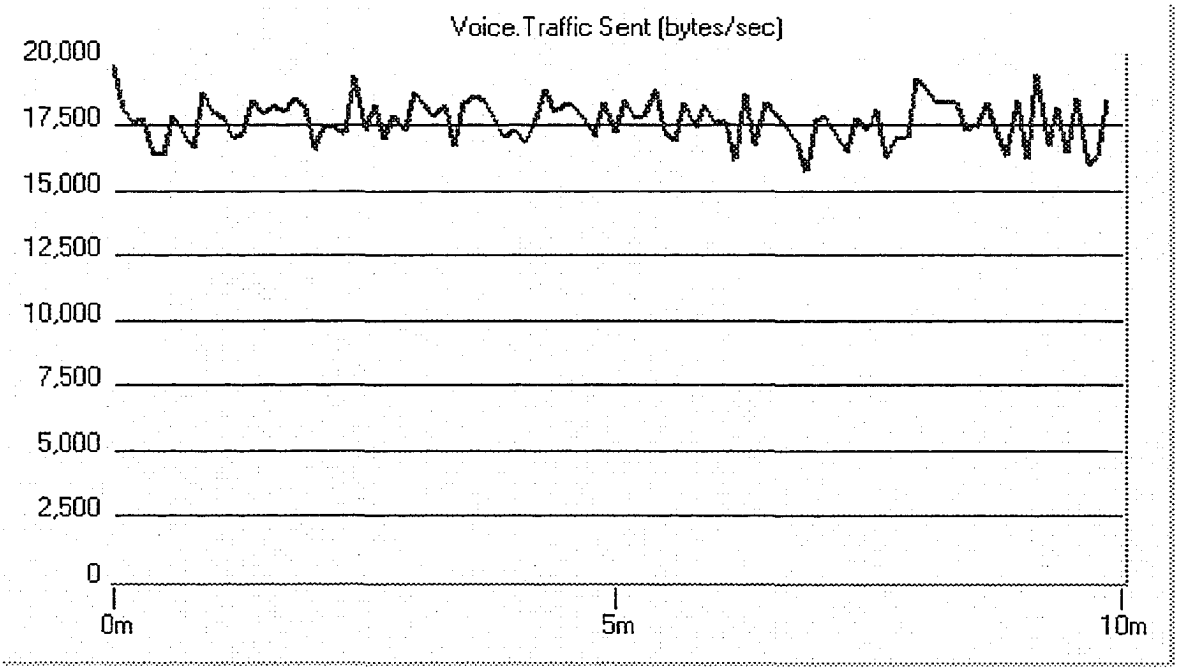


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graphical representation for voice traffic received:



- Graphical representation for voice traffic sent:

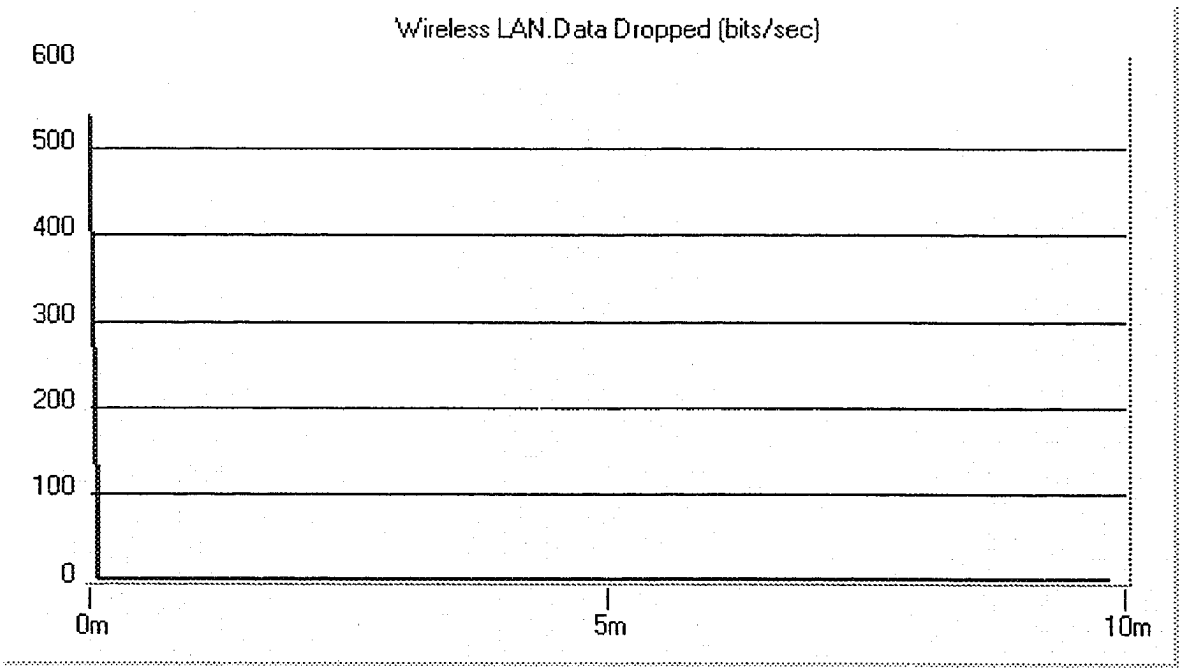


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- WLAN parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	5	536	0
Wireless LAN Delay (sec)	0.00042	0.00729	0.00026
Wireless LAN Load (bits/sec)	534,977	596,824	474,037
Wireless LAN Media Access Delay (sec)	0.00034	0.00684	0.00020
Wireless LAN Throughput (bits/sec)	534,972	596,288	474,037

- Graphical representation for voice packet drop rate:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.2. **For Frequency Hopping** the simulation results are:

- Voice quality parameters are:

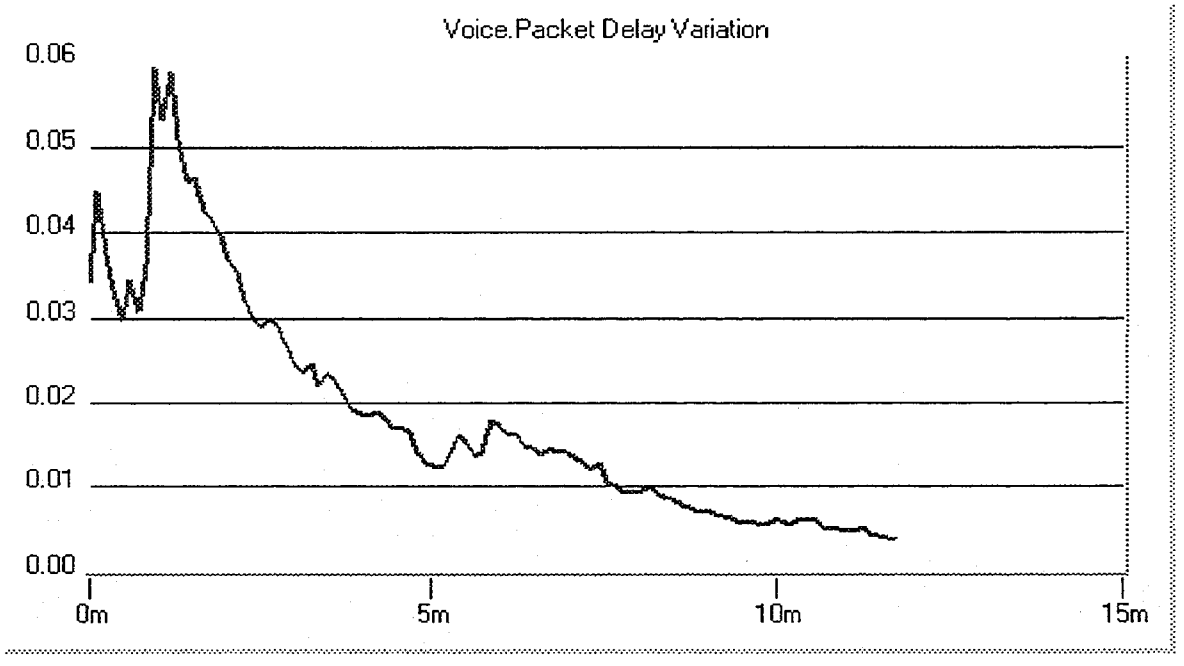
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.0183	0.0592	0.0035
Voice Packet End-to-End Delay (sec)	0.021	0.309	0.001
Voice Traffic Received (bytes/sec)	15,186	18,461	12,046
Voice Traffic Sent (bytes/sec)	15,198	19,187	12,044

We can notice that by using Frequency Hopping as a physical layer transmission protocol, and keeping all other parameters of the network unchanged, the parameters measuring voice quality: packet delay and delay variation, are in this case sensibly higher than in case of Direct Sequencing, but still below the acceptable threshold.

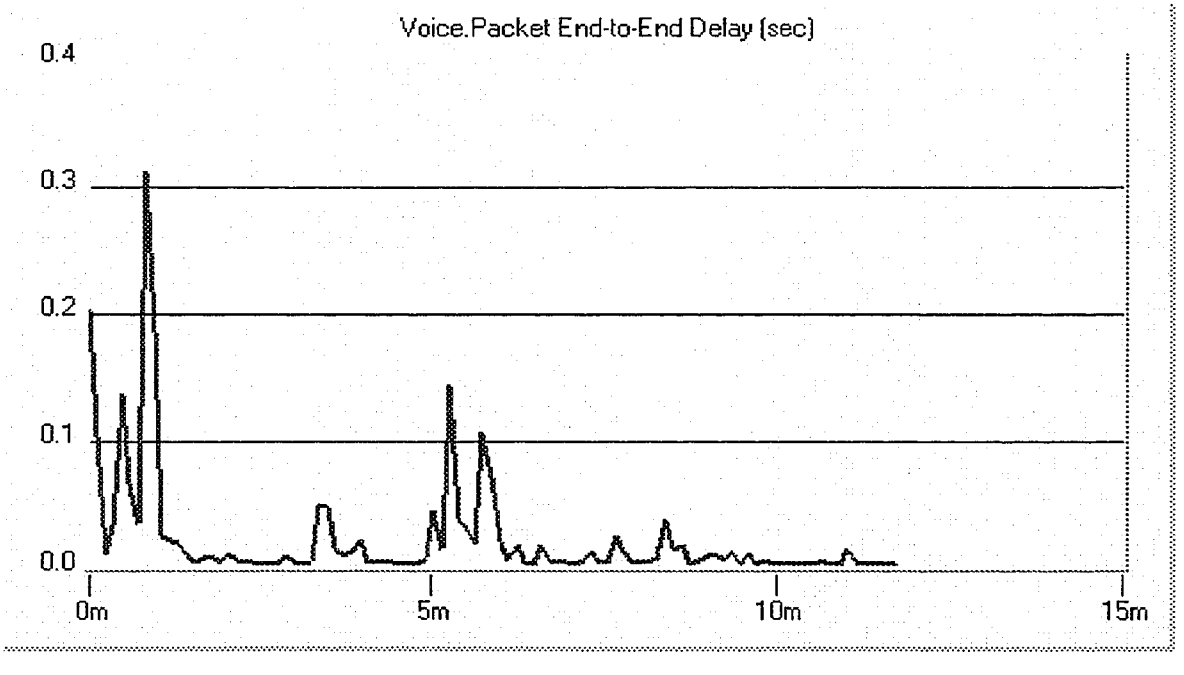
The graphical representation of the studied parameters are:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation for Voice Packet delay variation:

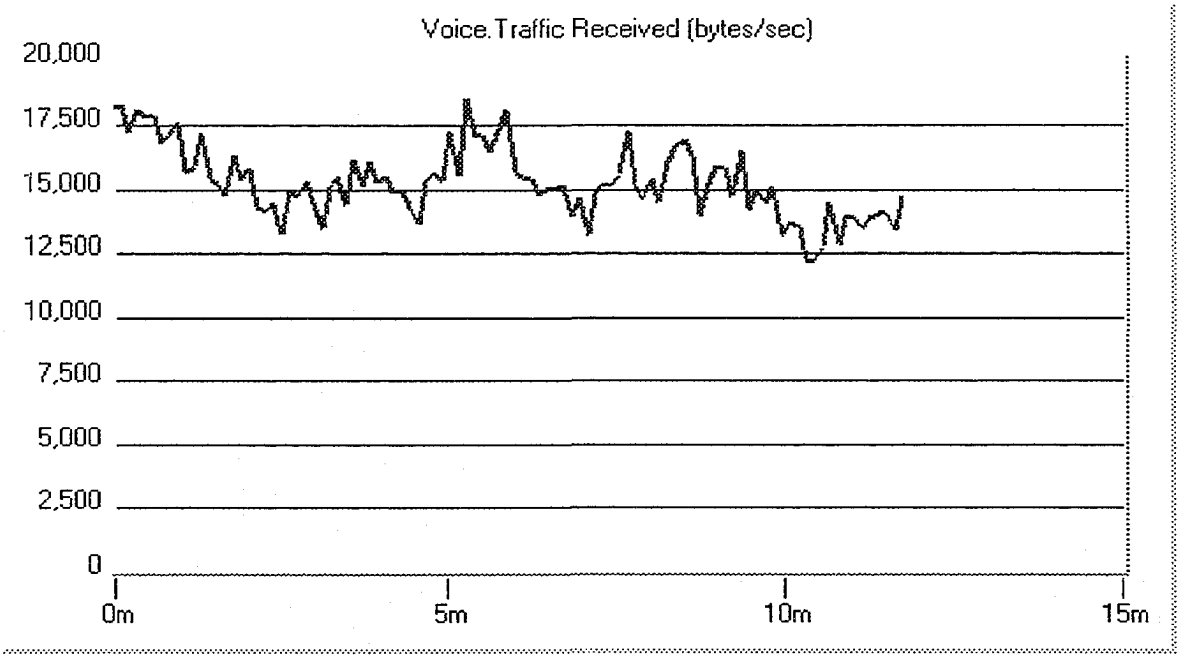


Graph representation of Voice Packet End-to-End Delay:

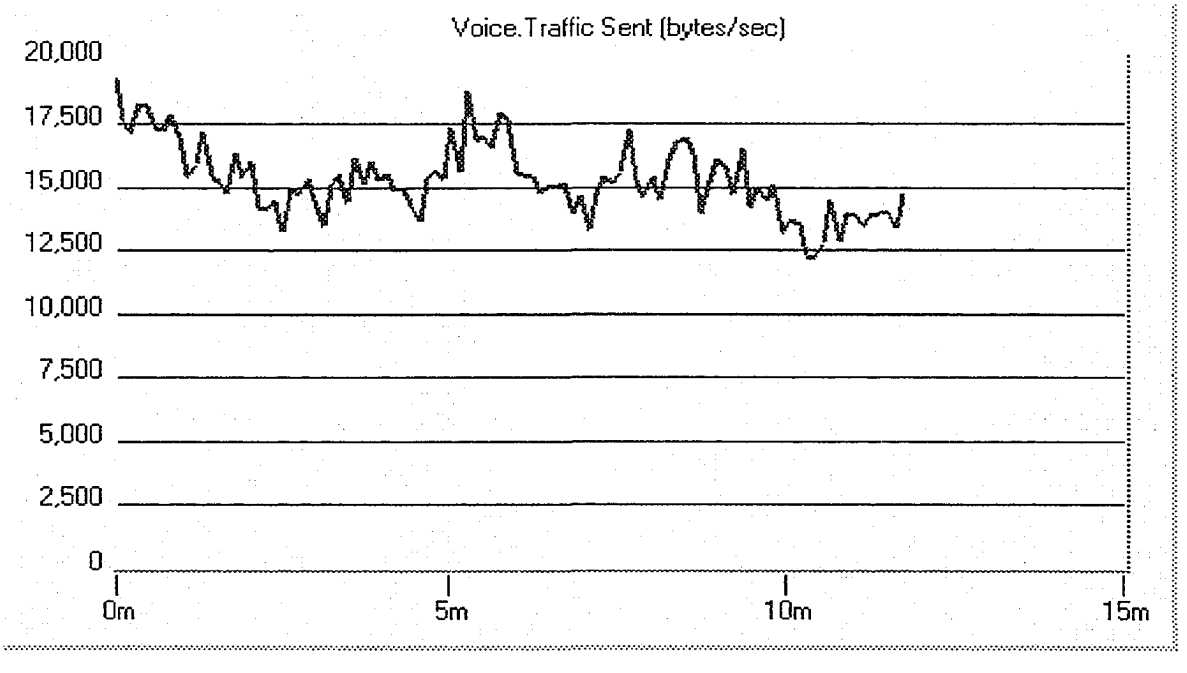


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Graph representation of Voice Traffic Received:



Graph representation of Voice Traffic Sent:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

WLAN parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	272	4,011	0
Wireless LAN Delay (sec)	0.021	0.308	0.000
Wireless LAN Load (bits/sec)	461,985	580,531	366,151
Wireless LAN Media Access Delay (sec)	0.020	0.303	0.000
Wireless LAN Throughput (bits/sec)	461,723	561,218	366,151

- Comparison table for DS vs. FH:

	Direct Sequence	Frequency Hopping
Statistic	Average	Average
Voice Packet Delay Variation	0.00006	0.0183
Voice Packet End-to-End Delay (sec)	0.00083	0.021
Voice Traffic Received (bytes/sec)	17,596	15,186
Voice Traffic Sent (bytes/sec)	17,599	15,198
Wireless LAN Data Dropped (bits/sec)	5	272

Comparing the two simulation results, we can notice that:

- The voice quality parameters, delay and delay variation are sensibly higher in case of frequency hopping.

- The traffic in both cases is symmetrical and close to each other in value.

- The packet-dropping rate is sensibly higher in case of frequency hopping.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As physical layer protocol, Direct Sequencing outperforms Frequency Hopping in voice transmission over wireless networks, by providing a faster connection accommodating more simultaneous users, with acceptable voice quality transmission.

2nd Scenario

In this scenario we study which voice codec is more performant.

Having this goal in mind, we set up a Wireless LAN with the following characteristics:

- 802.11b at 11Mbps
- ad-hoc network without Access Point
- 32 fix wireless workstations
- Physical Layer protocol is Direct Sequencing
- Calling scheme is as follows: workstation-1 calls workstation-2, workstation 3 calls workstation-4, ..., workstation31 calls workstation32; the voice traffic is bidirectional.

We study which voice codec is more performant:

- G.711, already widely used in telecom networks, or
- newly designed G.729.

After 10 minutes of network traffic simulation, the result for the significant parameters are:

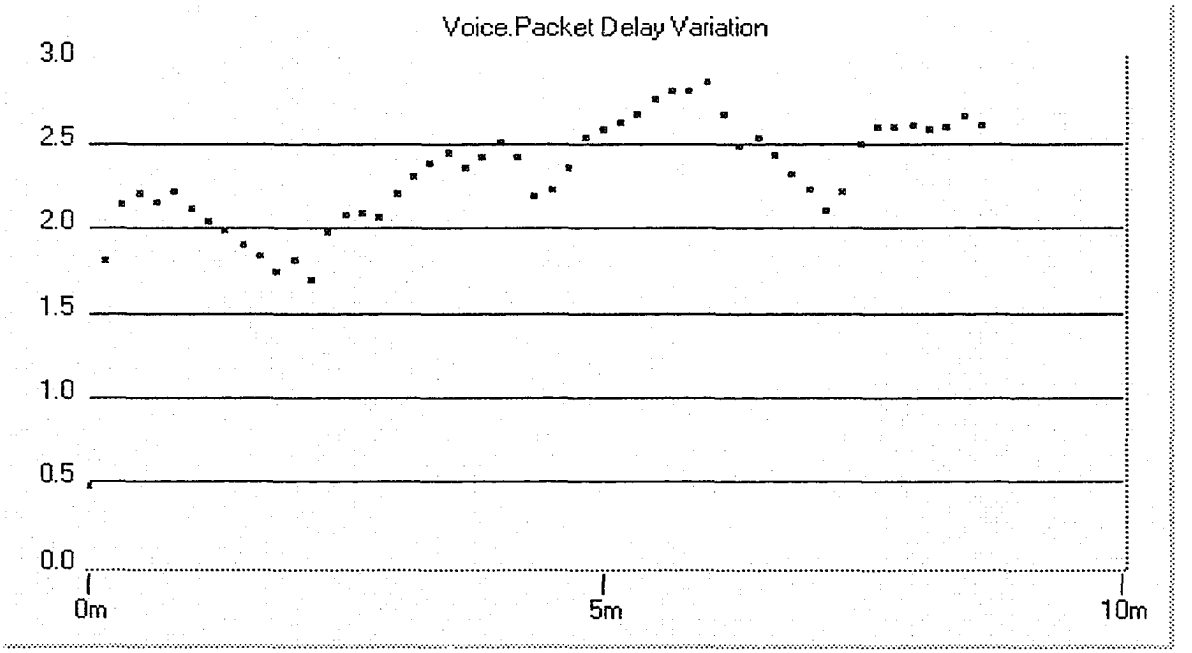
1. **For use of codec G.711** the simulation results are:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	2.28	2.85	0.47
Voice Packet End-to-End Delay (sec)	3.99	5.30	1.68
Voice Traffic Received (bytes/sec)	83,267	89,126	80,208
Voice Traffic Sent (bytes/sec)	120,761	143,171	104,349

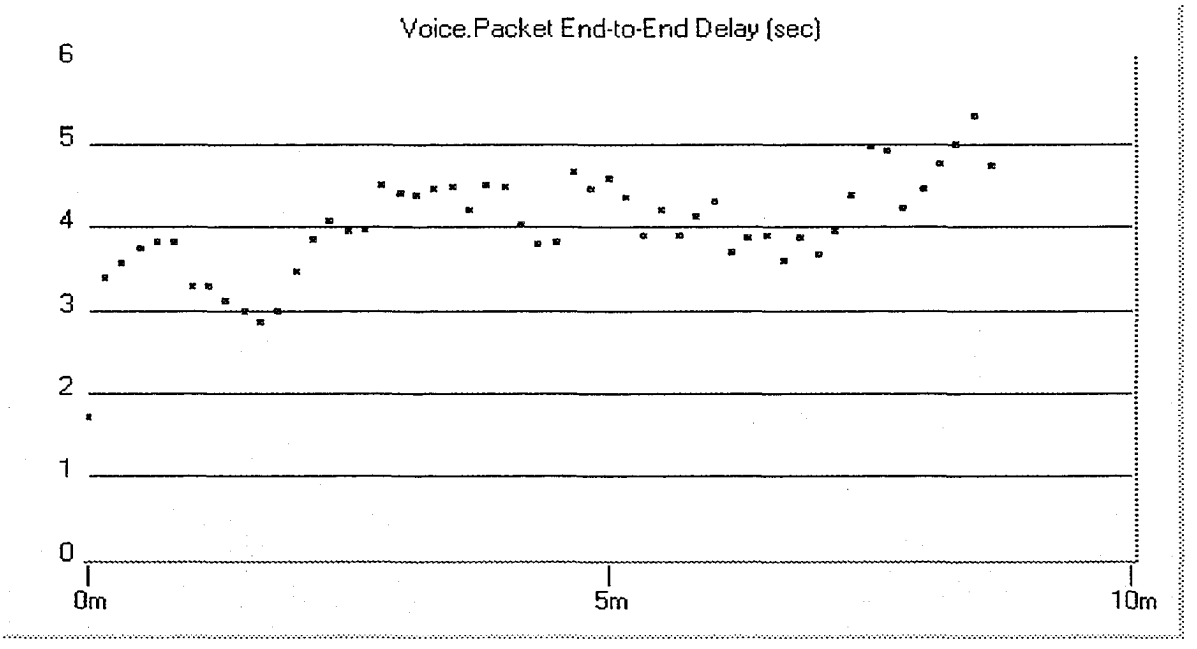
The graphical representation of the studied parameters are:

- Graph representation of Voice Packet delay variation:



- Graph representation of Voice Packet End-to-End Delay:

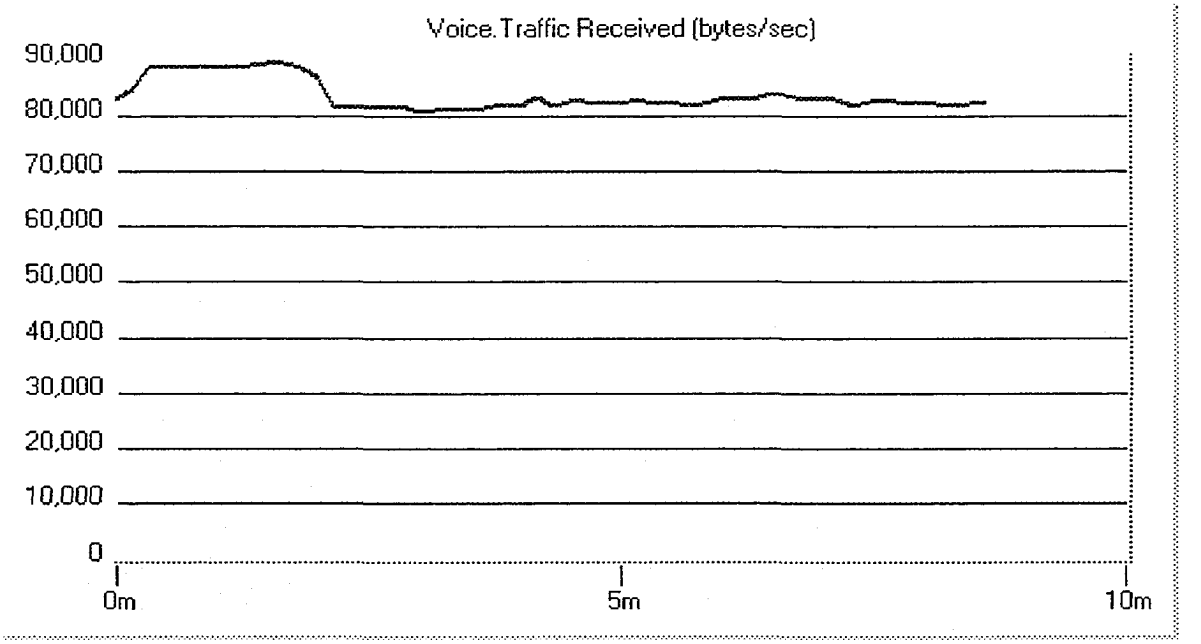
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



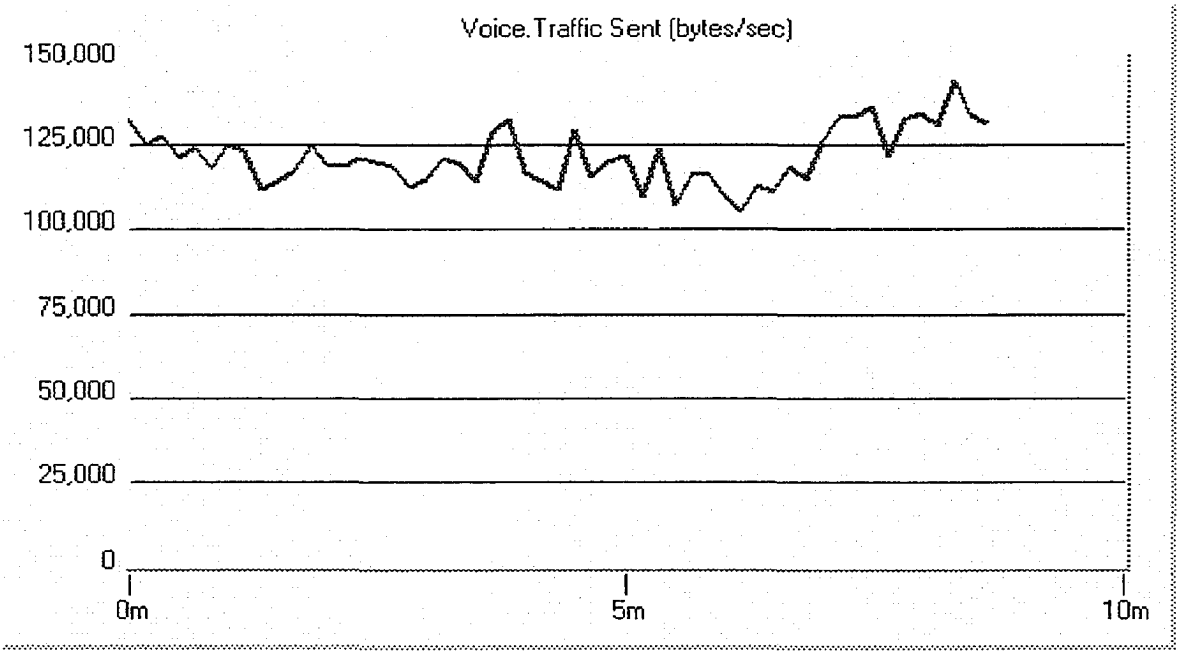
Note. The graph is not a continuous line because we decreased the sampling rate, in order to save simulation time.

- Graph representation of Voice Traffic Received:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- Graph representation of Voice Traffic Sent:



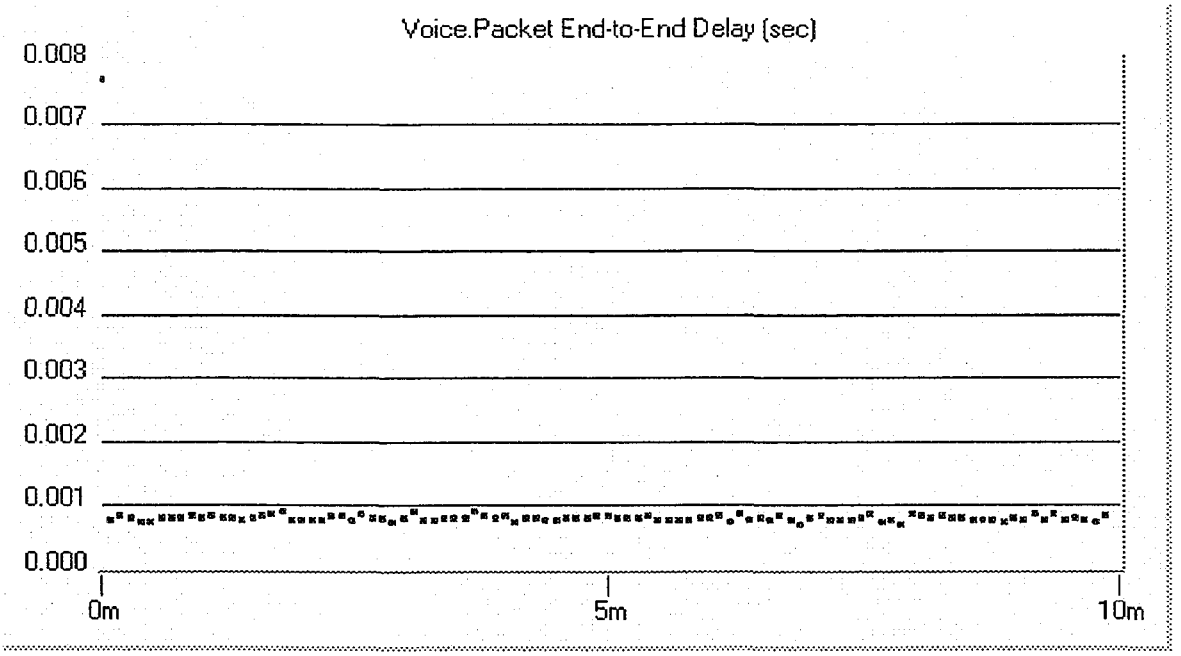
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. For use of codec G.729 the simulation results are:

- Voice quality parameters are:

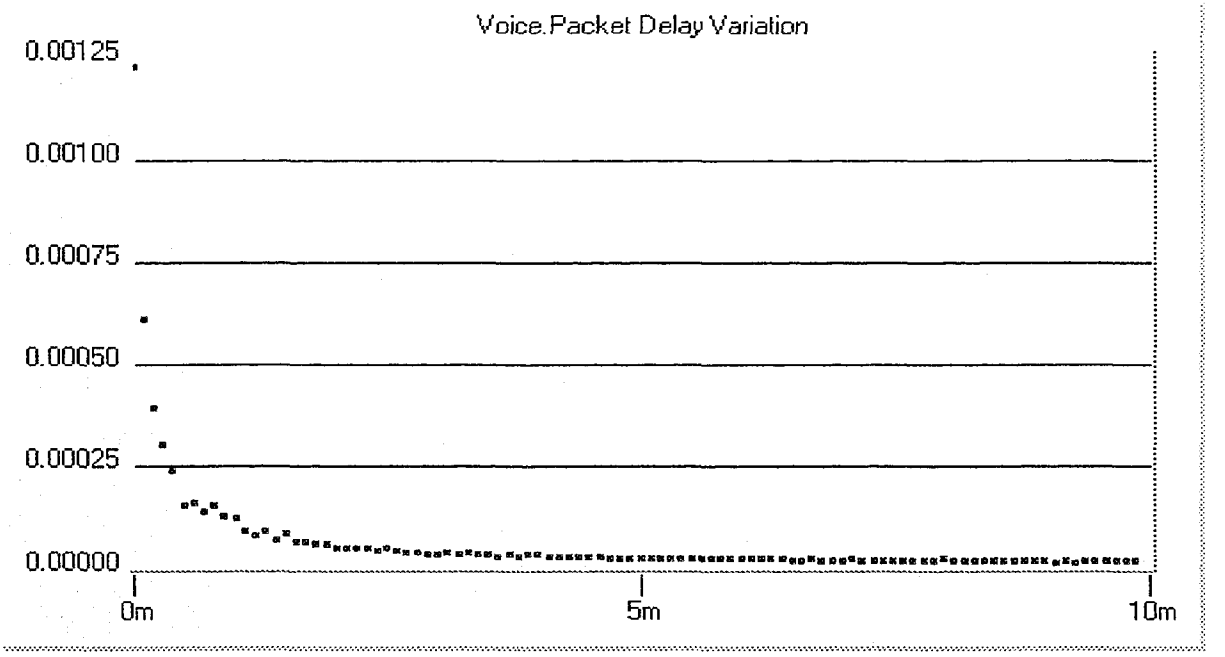
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.00006	0.00122	0.00001
Voice Packet End-to-End Delay (sec)	0.00083	0.00768	0.00067
Voice Traffic Received (bytes/sec)	17,596	19,480	15,592
Voice Traffic Sent (bytes/sec)	17,599	19,741	15,593

- Graphical representation for voice packet delay:



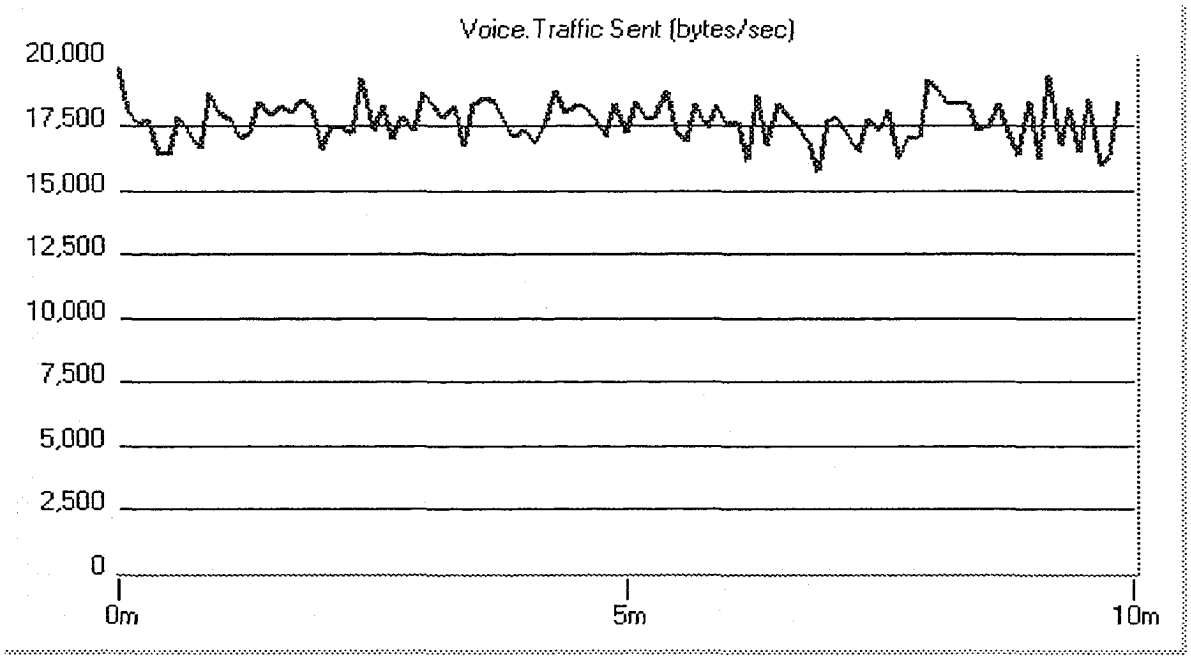
- Graphical representation for voice packet delay variation:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- Graphical representation for voice traffic sent:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

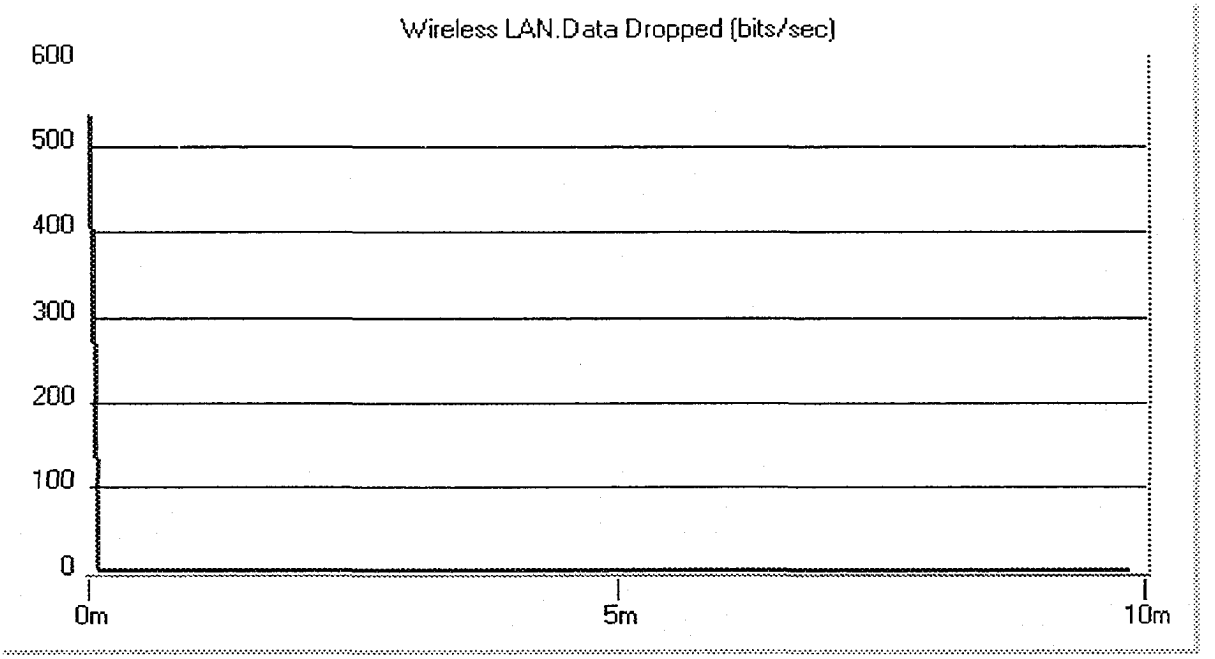


- WLAN parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	5	536	0
Wireless LAN Delay (sec)	0.00042	0.00729	0.00026
Wireless LAN Load (bits/sec)	534,977	596,824	474,037
Wireless LAN Media Access Delay (sec)	0.00034	0.00684	0.00020
Wireless LAN Throughput (bits/sec)	534,972	596,288	474,037

- Graphical representation for voice packet drop rate:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Comparison Table for G.711 vs. G.729:

In this table we compare the main voice quality parameters for the two type of simulations: one with codec G.711 and the second with codec G.729

	G.729	G.711
Statistic	Average(s)	Average(s)
Voice Packet Delay Variation	0.00006	2.28
Voice Packet End-to-End Delay (sec)	0.00083	3.99
Voice Traffic Received (bytes/sec)	17,596	83,267
Voice Traffic Sent (bytes/sec)	17,599	120,761

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Conclusion:

Comparing the two simulation results, for codec G.729 and G.711 we can notice that G.729 is, in our simulation, sensibly more efficient: when generating an equal number of simultaneous voice calls, G. 729 has a significantly lower delay and delay variation.

As a result of the simulations, we can conclude that codec G.729 outperforms G.711, and is much more suitable for packet switched voice transmission in wireless networks.

3th Scenario

In this scenario we study the impact of the workstation mobility on the performance of the voice transmission in a wireless network.

With this goal in mind, we set up a Wireless LAN with the following characteristics:

- 802.11b at 11Mbps
- ad-hoc network without Access Point
- Physical Layer protocol is Direct Sequencing
- calling scheme is as follows:
- workstation-1 calls workstation-2, workstation 3 calls workstation-4, ..., workstation31 calls workstation32, with bidirectional voice traffic.
- voice codec used is first G.711 and then G.729

We study the impact of the workstation mobility on the performance of the voice transmission:

We setup a wireless network with 32 fixed wireless workstations

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and we compare the results of the 2 simulations.

For each mobile workstation a path is chosen and defined randomly, inside the transmission area. Because of limited distances (the area is 1000ftx1000ft), the speed is very low and the mobile user is alternating 40 seconds of walking speed with 20 seconds of resting (0 speed).

Simulation results:

1. **For fixed wireless workstations**:

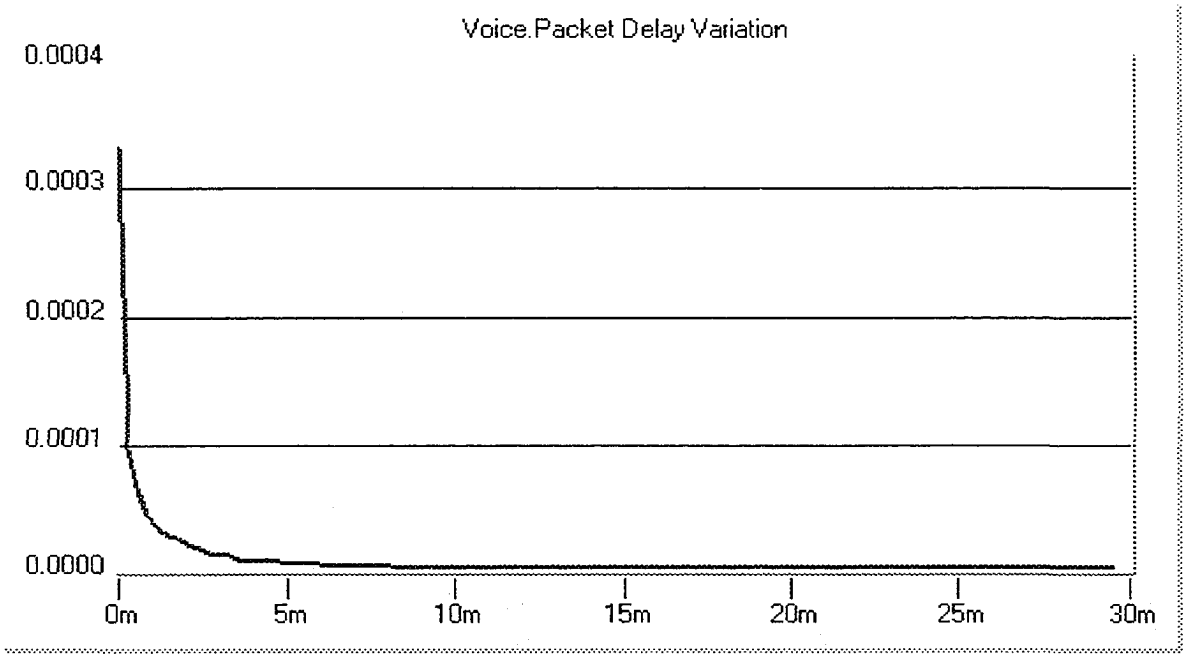
1.1 The voice quality parameters are:

Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.000008	0.000331	0.000000
Voice Packet End-to-End Delay (sec)	0.00064	0.00200	0.00051
Voice Traffic Received (bytes/sec)	12,626	17,284	7,499
Voice Traffic Sent (bytes/sec)	12,629	17,284	7,499

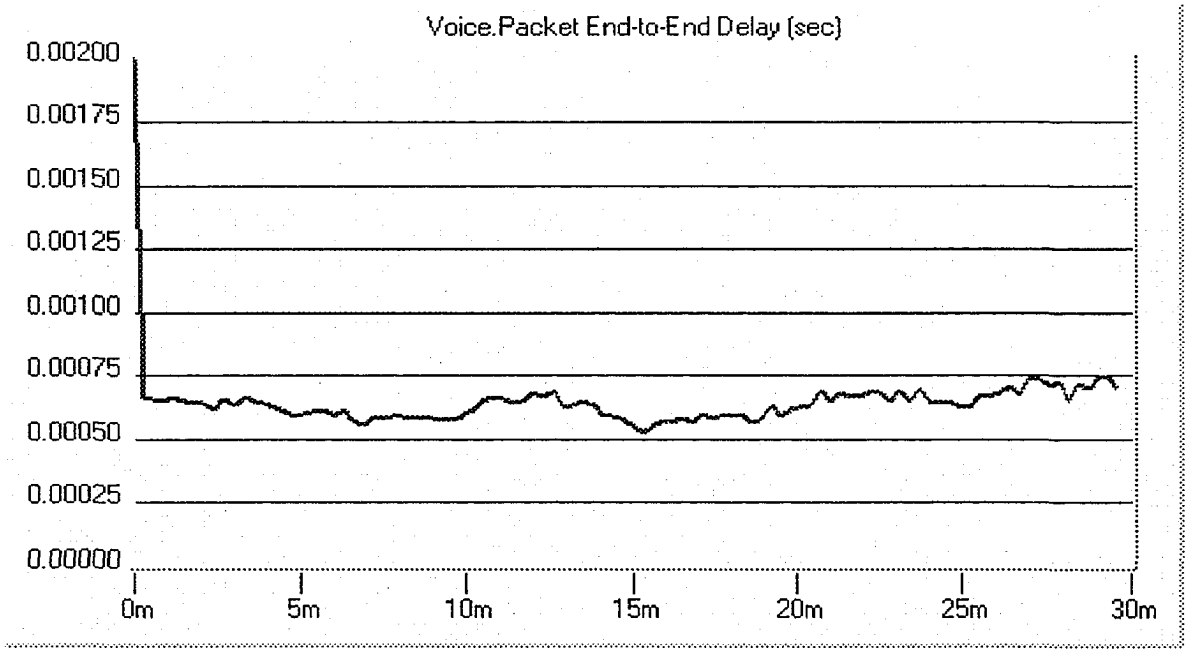
The graphical representation of the studied parameters are:

- Graph representation of Voice Packet delay variation:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

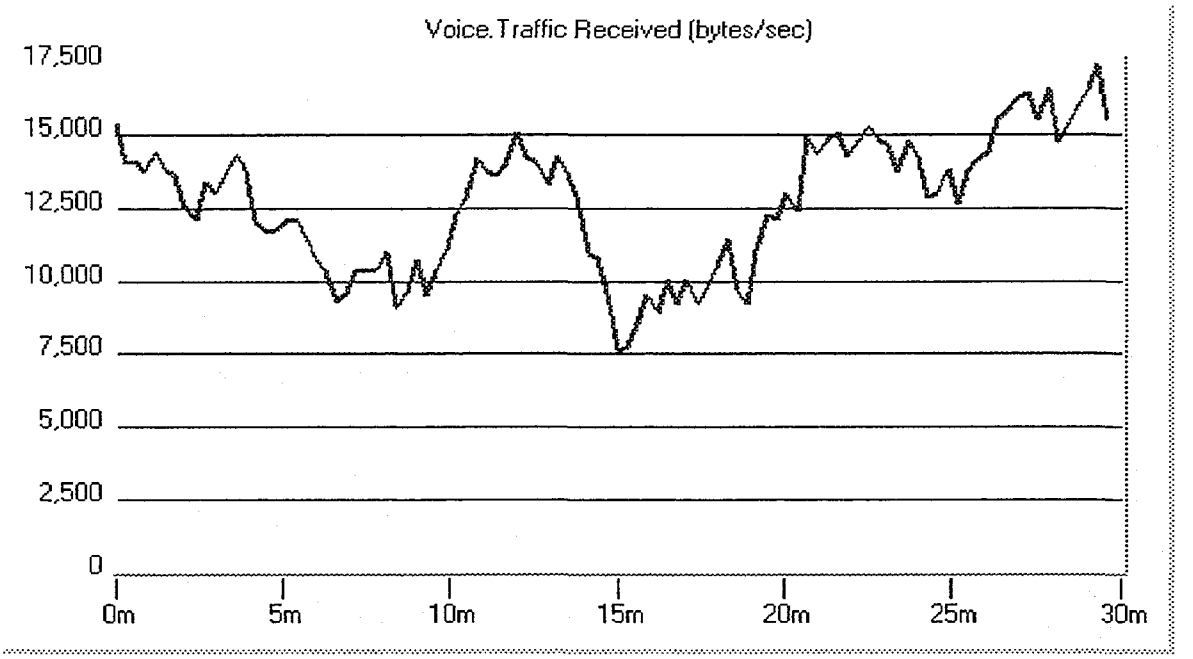


- Graph representation of Voice Packet End-to-End Delay:

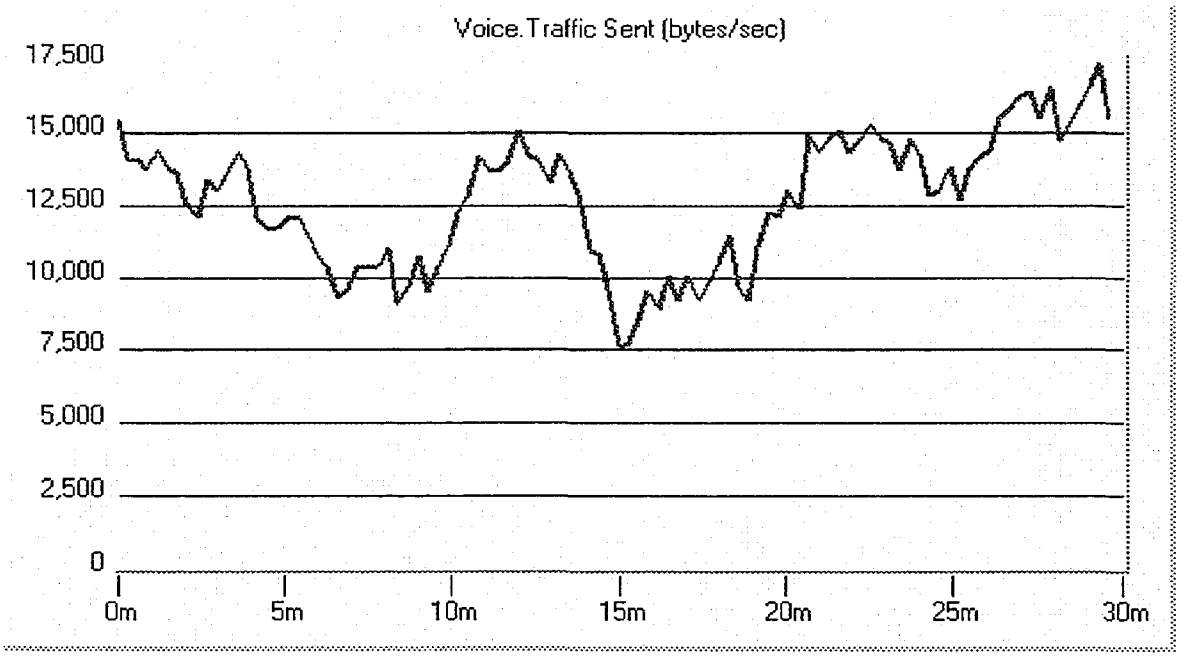


- Graph representation of Voice Traffic Received:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- Graph representation of Voice Traffic Sent:



2. **For mobile wireless workstations** the simulation results are:

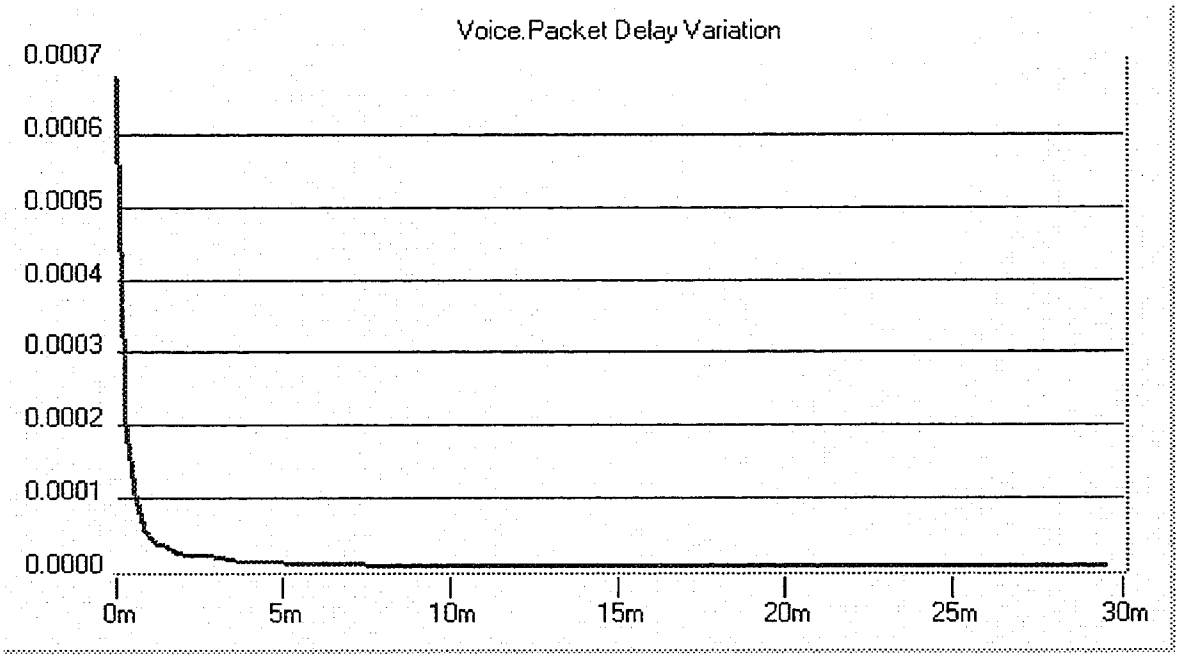
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- The voice quality parameters are:

Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.000013	0.000679	0.000000
Voice Packet End-to-End Delay (sec)	0.00067	0.00302	0.00057
Voice Traffic Received (bytes/sec)	13,691	17,741	10,456
Voice Traffic Sent (bytes/sec)	13,695	17,741	10,456

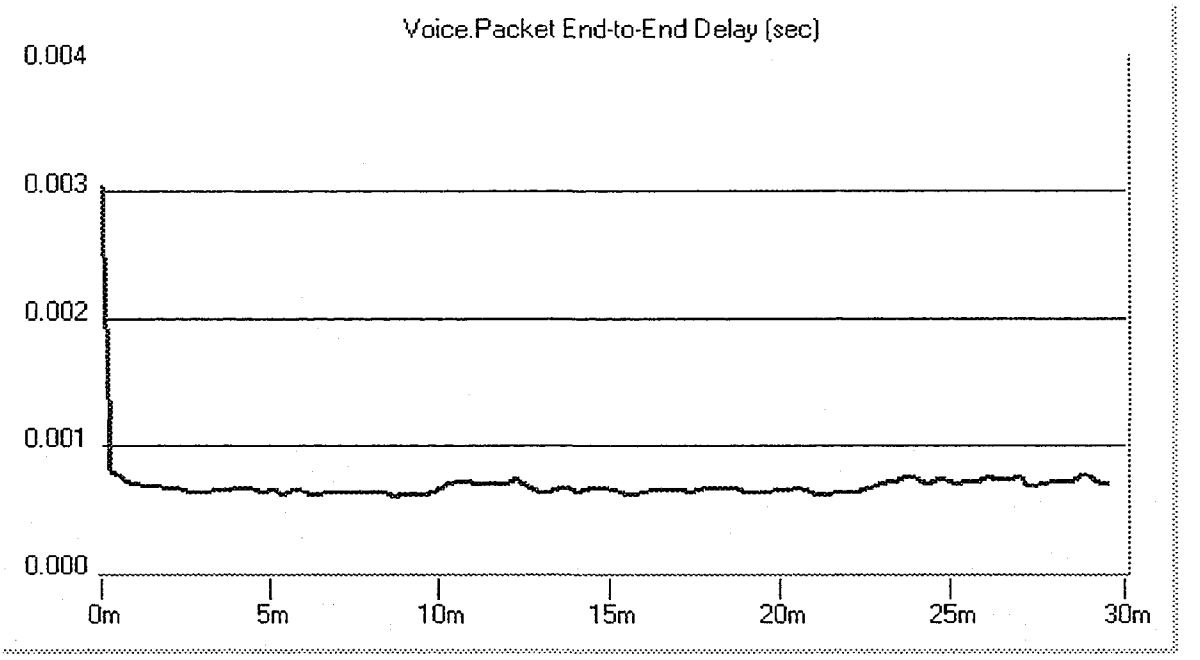
The graphical representation of the studied parameters are:

- Graph representation of Voice Packet delay variation:

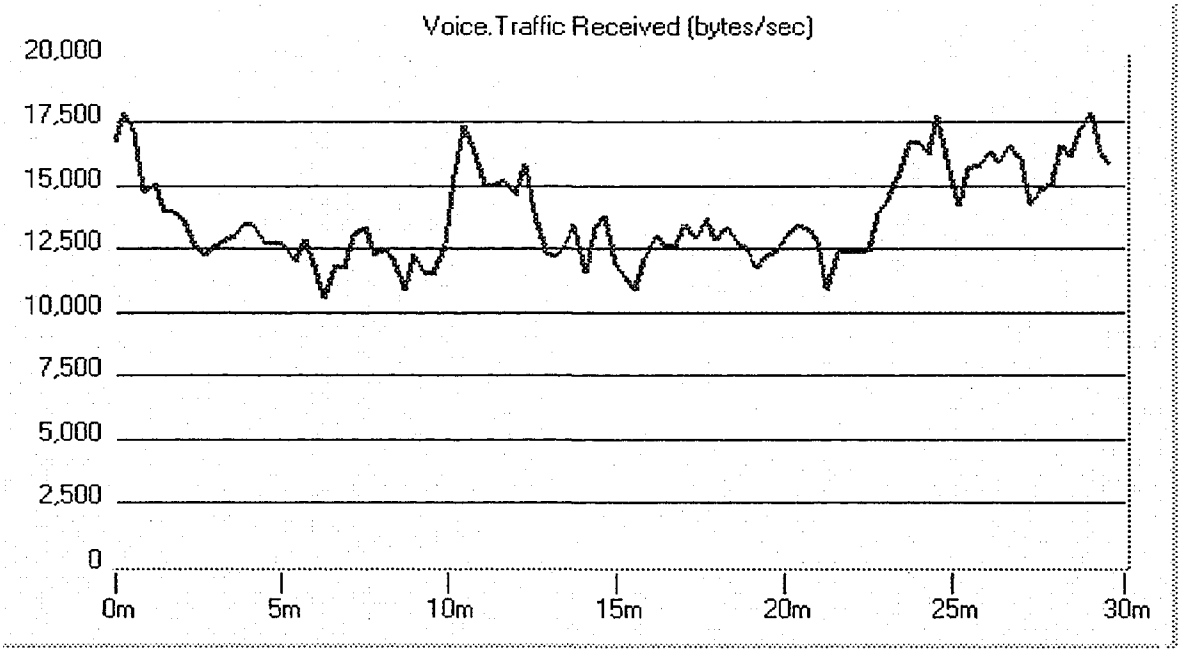


- Graph representation of Voice Packet End-to-End Delay:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

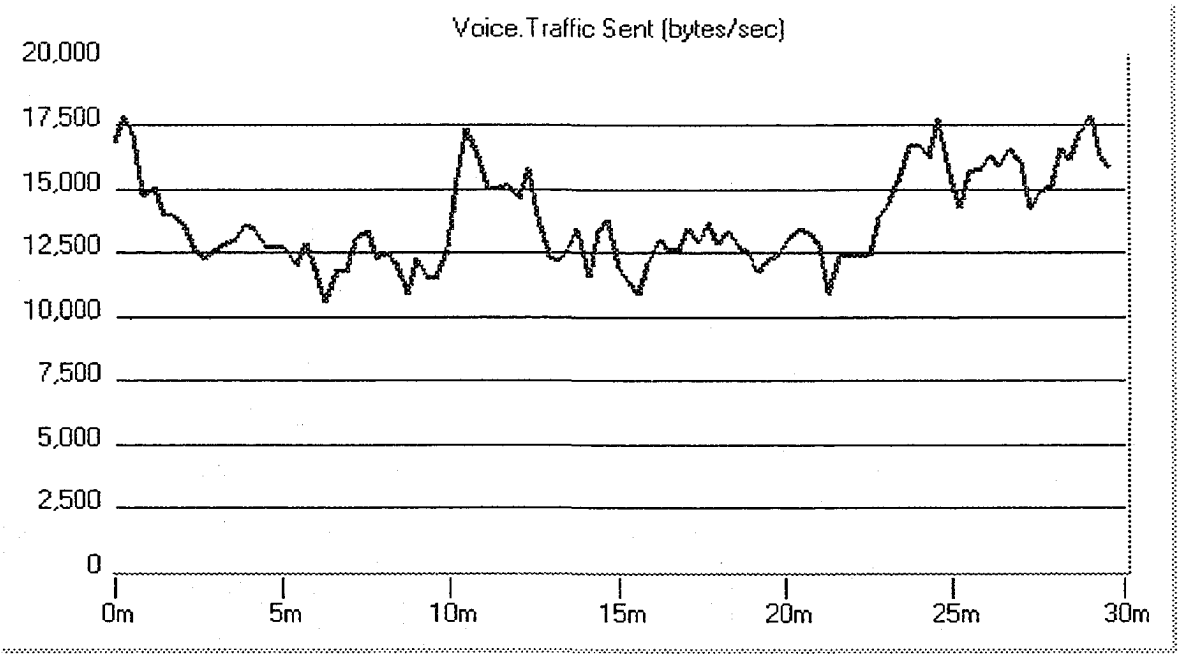


- Graph representation of Voice Traffic Received:



- Graph representation of Voice Traffic Sent:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Comparison table for fixed vs. mobile:

The main voice quality parameters are compared bellow:

	Fixwd wkst.	Mobile wkst.
Statistic	Average	Average
Voice Packet Delay Variation	0.000008	0.000013
Voice Packet End-to-End Delay (sec)	0.00064	0.00067
Voice Traffic Received (bytes/sec)	12,626	13,691
Voice Traffic Sent (bytes/sec)	12,629	13,695
Wireless LAN Data Dropped (bits/sec)	2	2

Conclusion for mobile vs. fixed stations:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Comparing the two sub-scenarios, we can notice that:

- Both, delay and delay variation are slightly higher in case of mobile workstations
- Traffic is symmetrical in both subscenarios and close in value to each other
- The packet drop rate is the same and very low.

We can conclude that mobility, when limited in speed and distance, has some minor negative impact on the performance of voice transmission.

4th Scenario

In this scenario we study the impact on the performance of voice transmission by the number of wireless workstations.

Having this goal in mind, we set up a Wireless LAN with the following characteristics:

- 802.11b at 11Mbps

- ad-hoc network without Access Point

- Physical Layer protocol is Direct Sequencing

- calling scheme is as follows: workstation-1 calls workstation-2, workstation 3 calls workstation-4, ..., etc. and traffic is bidirectional

- voice codec used is G.729

We study the impact on the performance of voice transmission by the number of wireless workstations. In subscenarios, we e simulate a wireless network with all parameters kept constant except the number of wireless workstations being increased as follows: consecutively 16, 32, 48, simultaneous voice calls are generated, and we compare the voice performance statistics.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Because of the increased number of computations as the number of workstations increases, we reduced simulation time to as low as 10 minutes and the number of collected statistics from 100 to as low as 50.

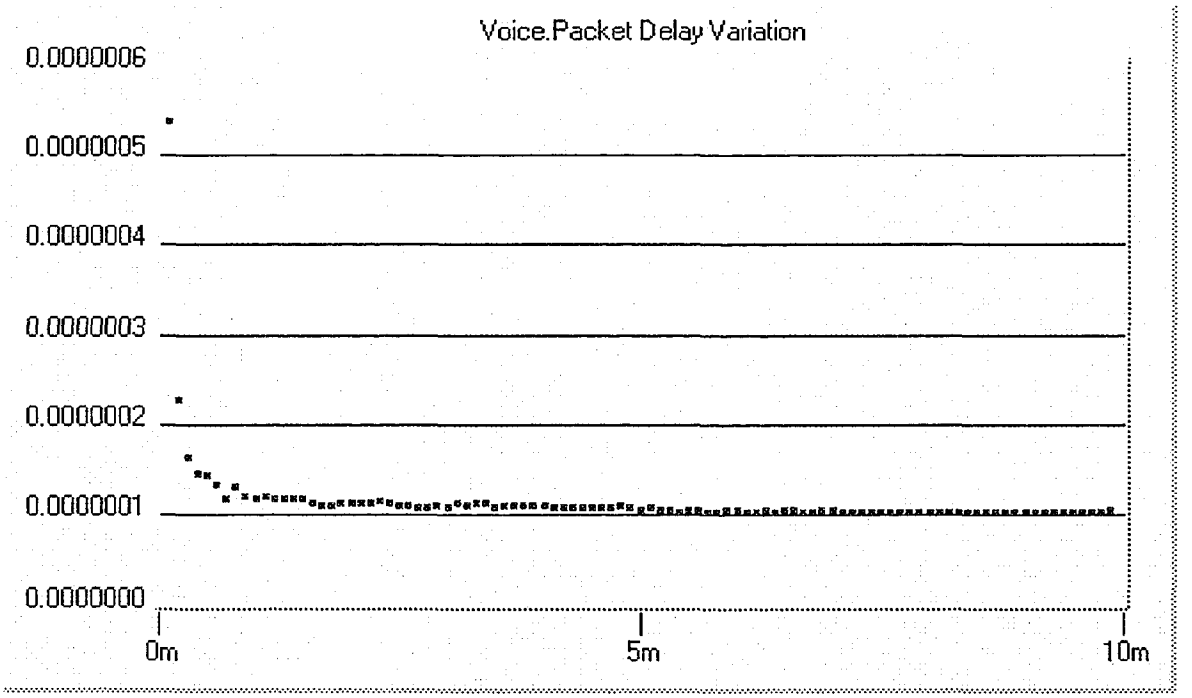
Simulation results:

Subscenario 4.1: For 16 fixed wireless workstations:

- Voice quality parameters are:

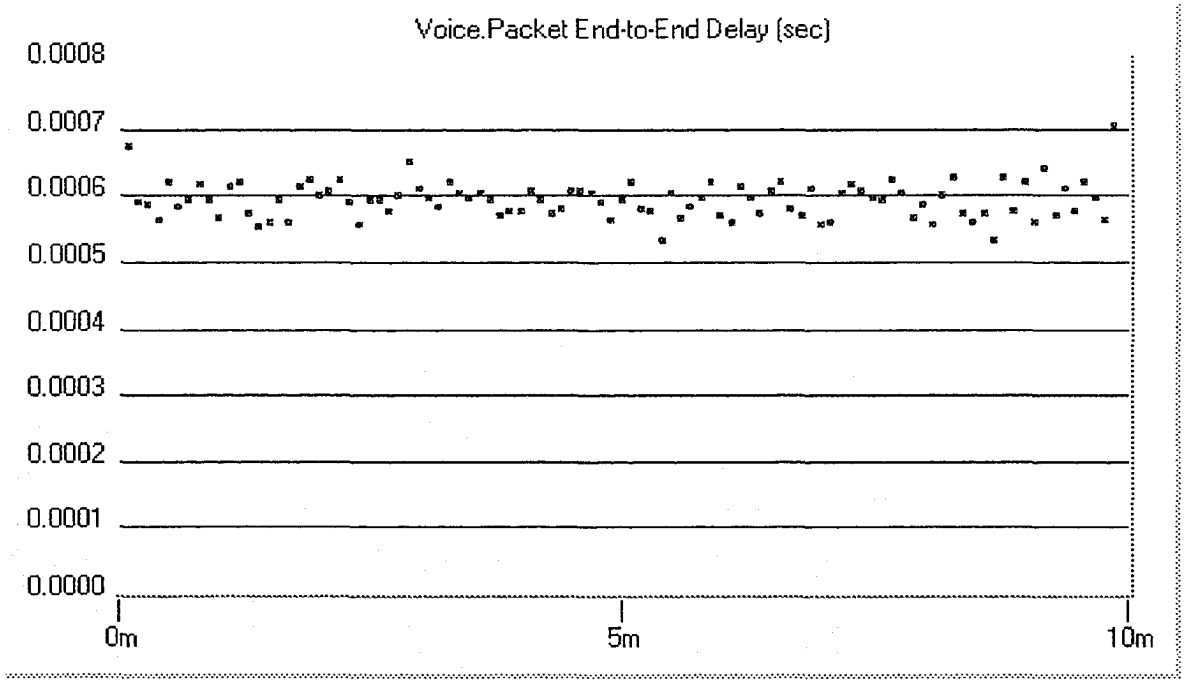
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.000000113	0.000000535	0.000000100
Voice Packet End-to-End Delay (sec)	0.000591	0.000703	0.000530
Voice Traffic Received (bytes/sec)	10,730	12,292	0
Voice Traffic Sent (bytes/sec)	10,731	12,293	0

- Graphical representation for voice packet delay variation:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

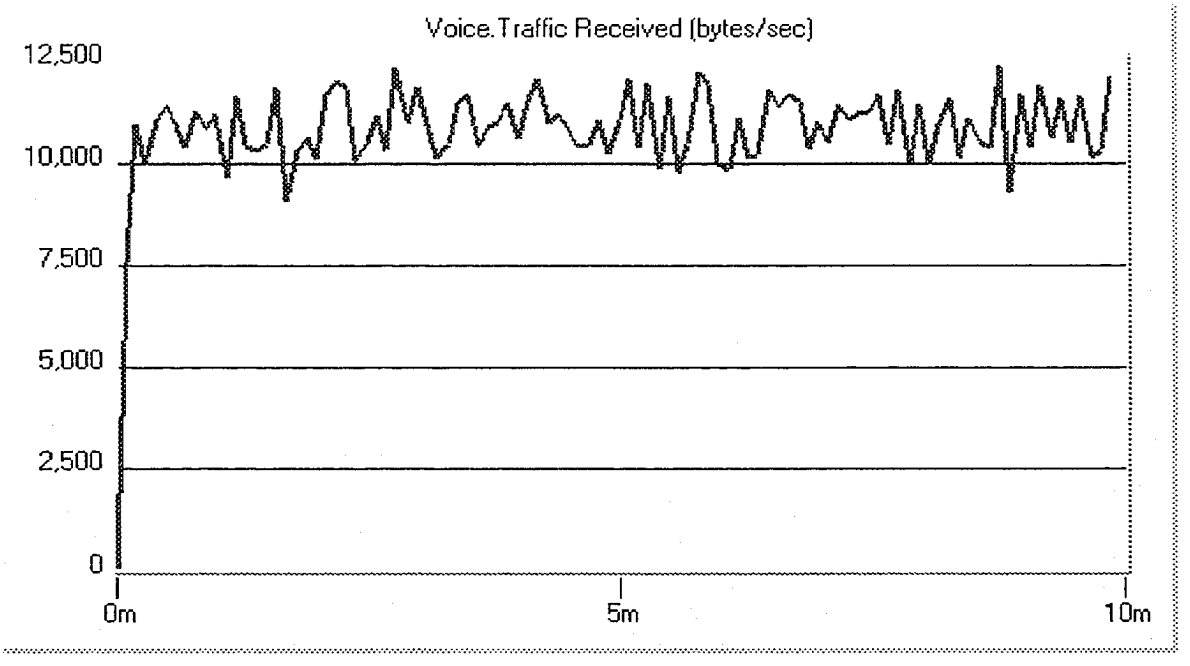
- Graphical representation for voice packet delay:



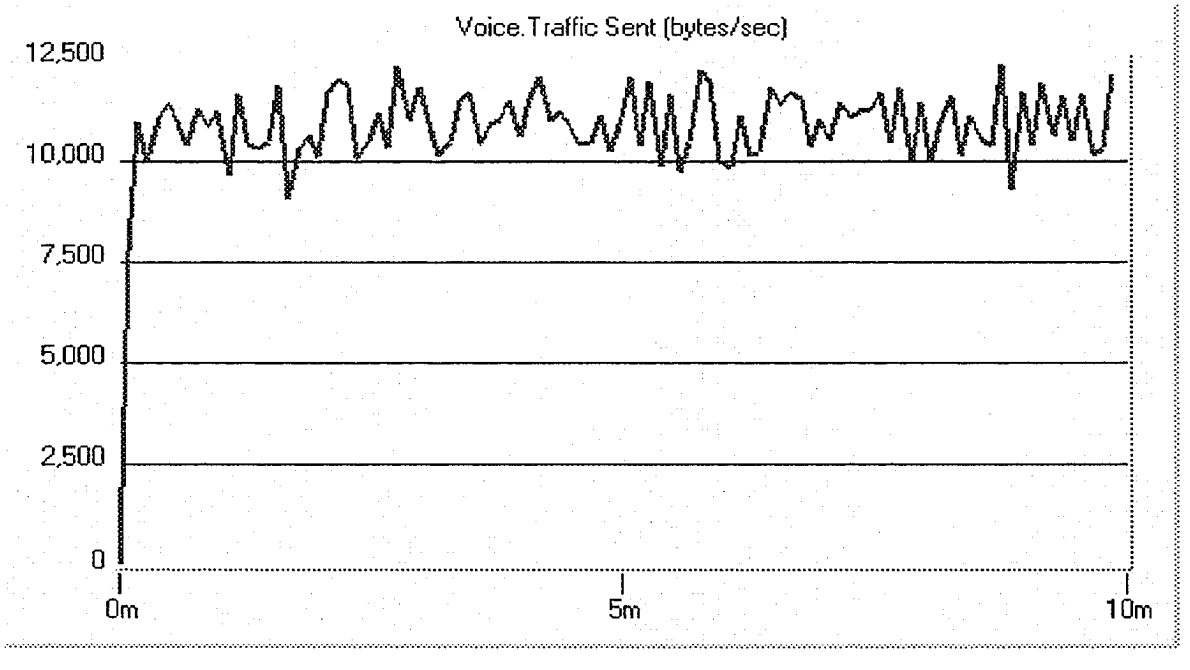
- Graphical representation for voice traffic received:

-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- Graphical representation for voice traffic sent:

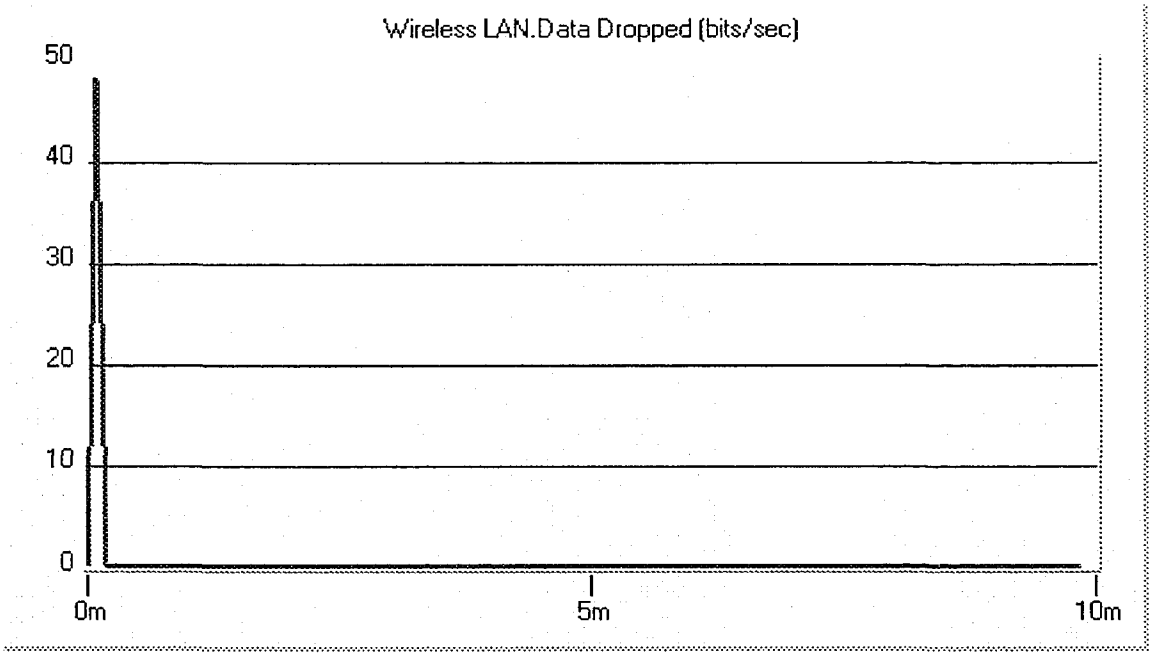


- WLAN traffic parameters:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	0.5	48.0	0.0
Wireless LAN Delay (sec)	0.000188	0.000302	0.000127
Wireless LAN Load (bits/sec)	326,215	373,717	0
Wireless LAN Media Access Delay (sec)	0.000134	0.000230	0.000077
Wireless LAN Throughput (bits/sec)	326,214	373,667	0

- Graphical representation for voice packet drop rate:



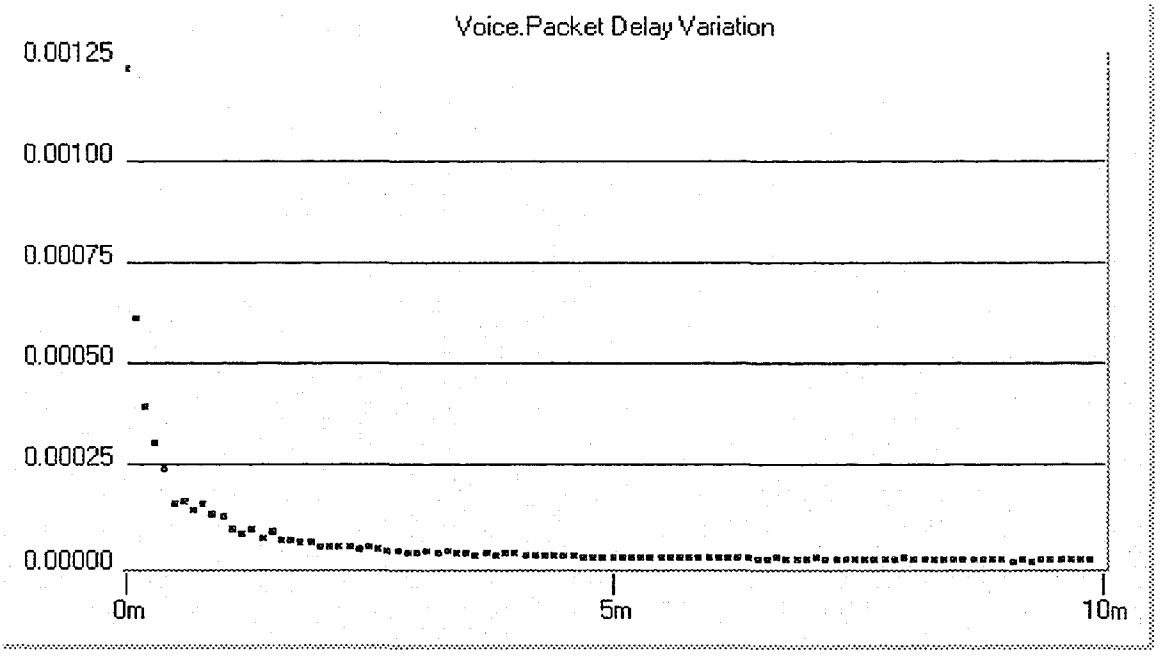
Subsenario 4.2: 32 fixed wireless workstations:

- Voice quality parameters are:

Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.00006	0.00122	0.00001
Voice Packet End-to-End Delay (sec)	0.00083	0.00768	0.00067
Voice Traffic Received (bytes/sec)	17,596	19,480	15,592
Voice Traffic Sent (bytes/sec)	17,599	19,741	15,593

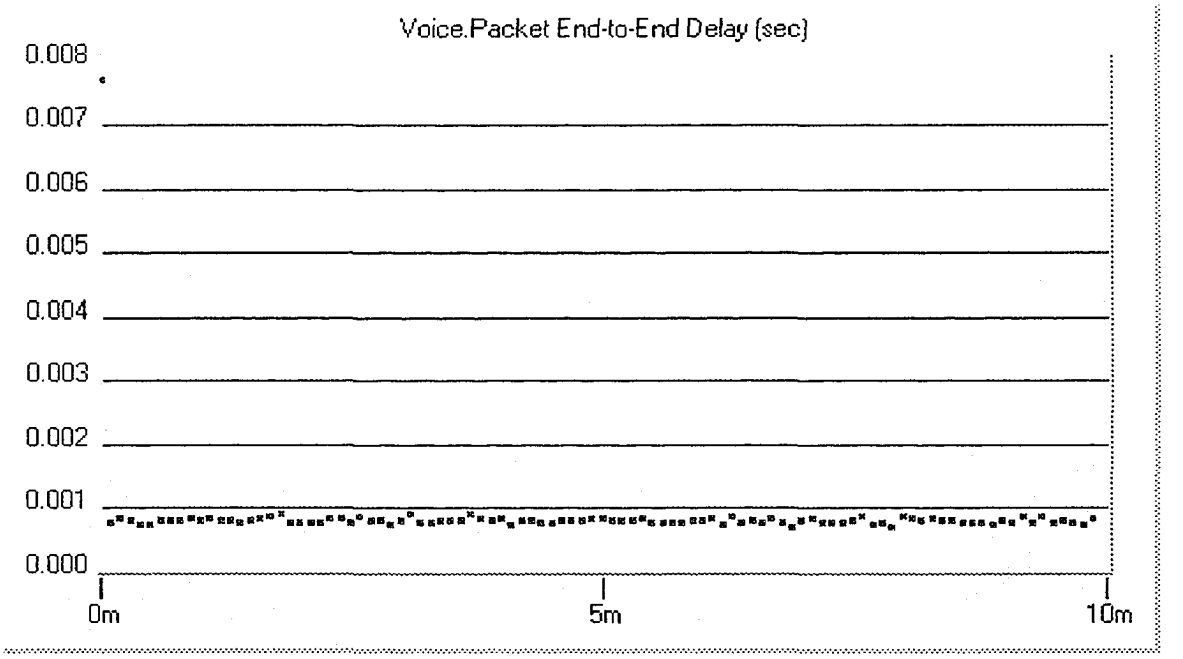
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graphical representation for voice packet delay variation:

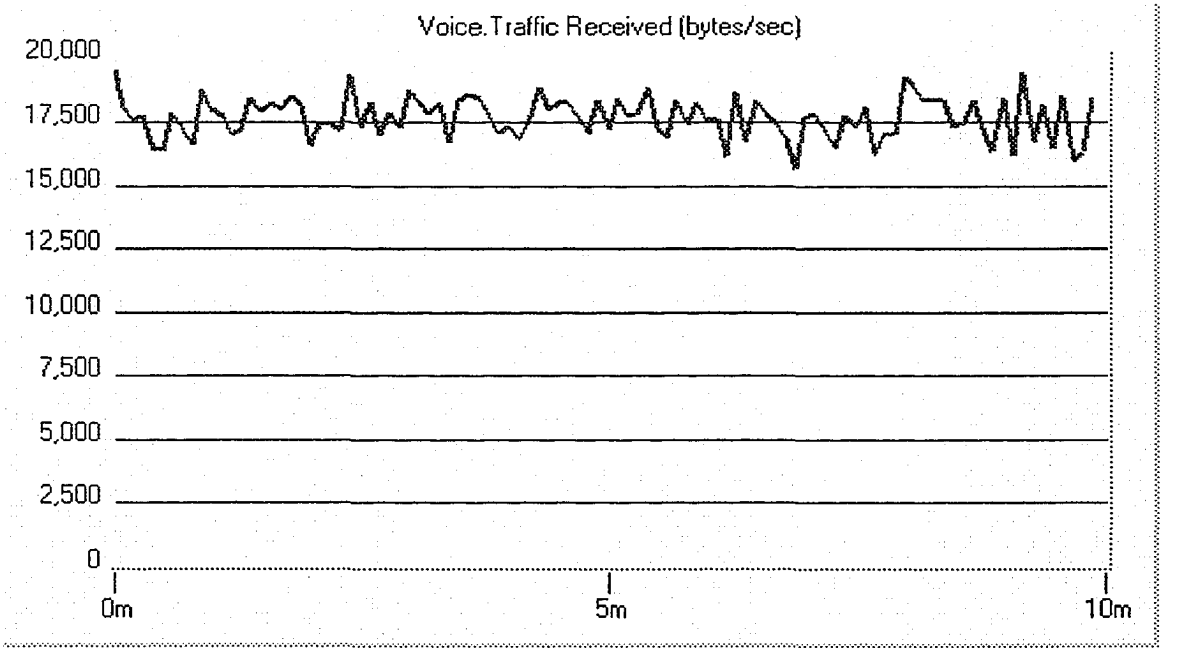


- Graphical representation for voice packet delay:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

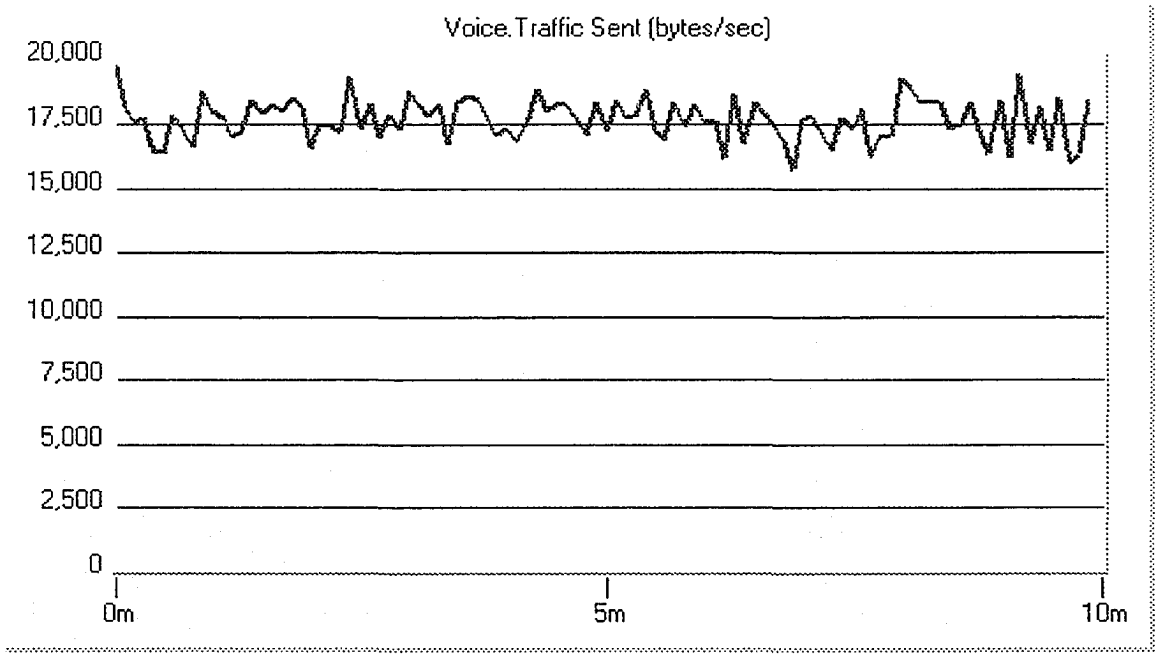


- Graphical representation for voice traffic received:



- Graphical representation for voice traffic sent:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

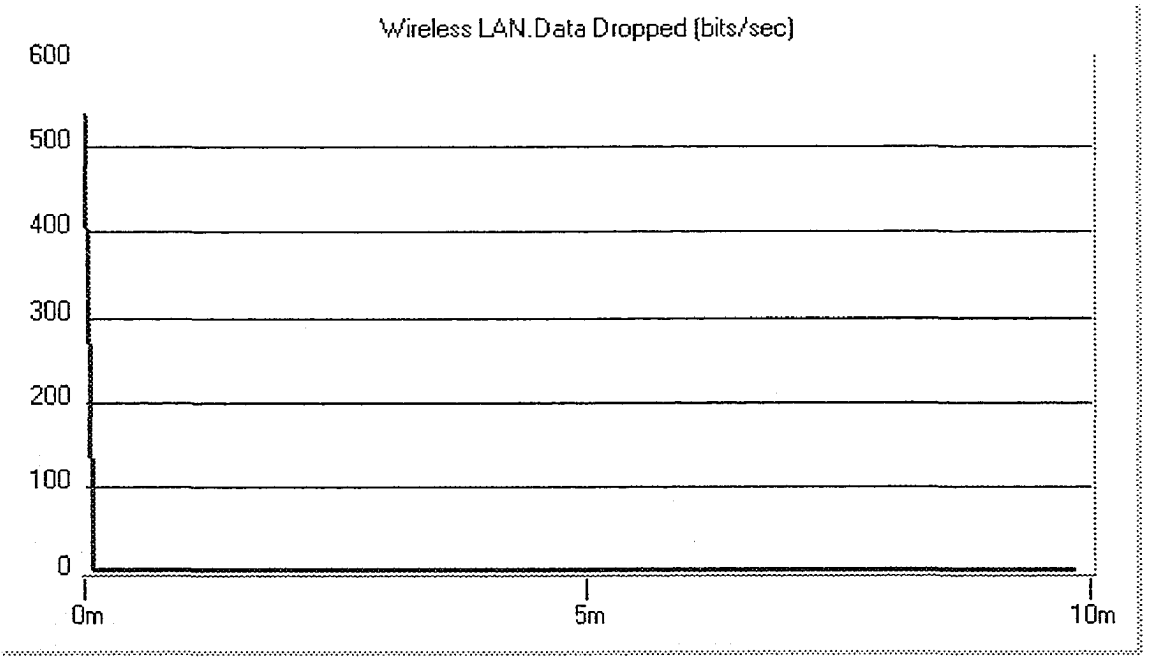


- WLAN parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	5	536	0
Wireless LAN Delay (sec)	0.00042	0.00729	0.00026
Wireless LAN Load (bits/sec)	534,977	596,824	474,037
Wireless LAN Media Access Delay (sec)	0.00034	0.00684	0.00020
Wireless LAN Throughput (bits/sec)	534,972	596,288	474,037

- Graphical representation for voice packet drop rate:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



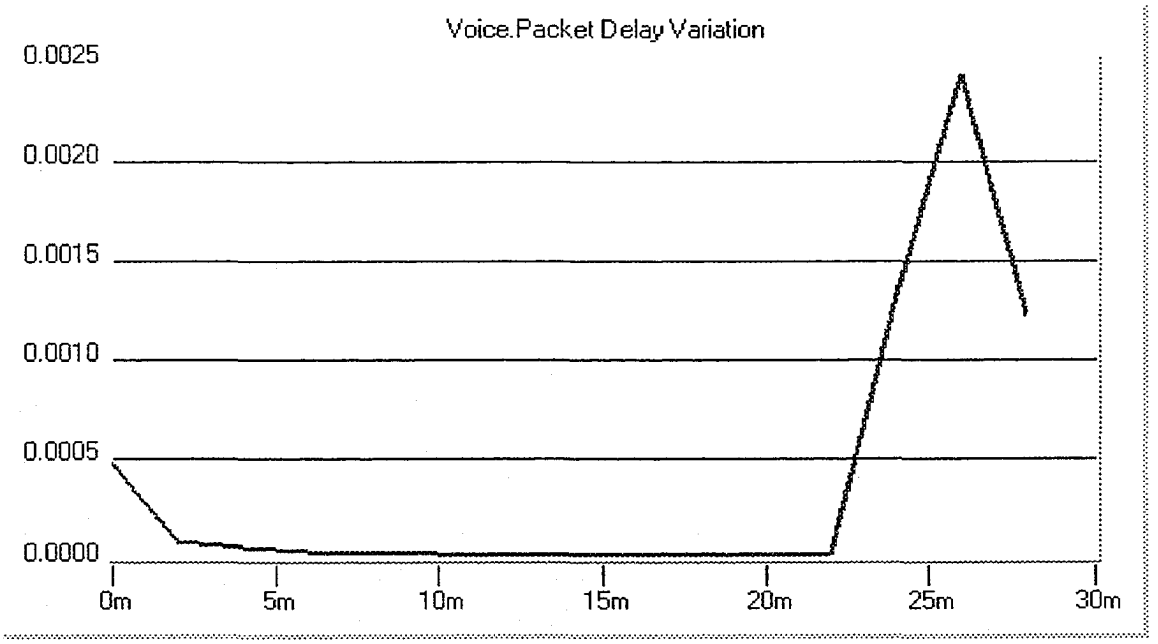
Subscenario 4.3 For 48 fixed wireless workstations:

-. The voice quality simulation results are:

Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.00038	0.00242	0.00000
Voice Packet End-to-End Delay (sec)	0.0024	0.0221	0.0006
Voice Traffic Received (bytes/sec)	18,811	25,338	13,031
Voice Traffic Sent (bytes/sec)	18,815	25,351	13,036

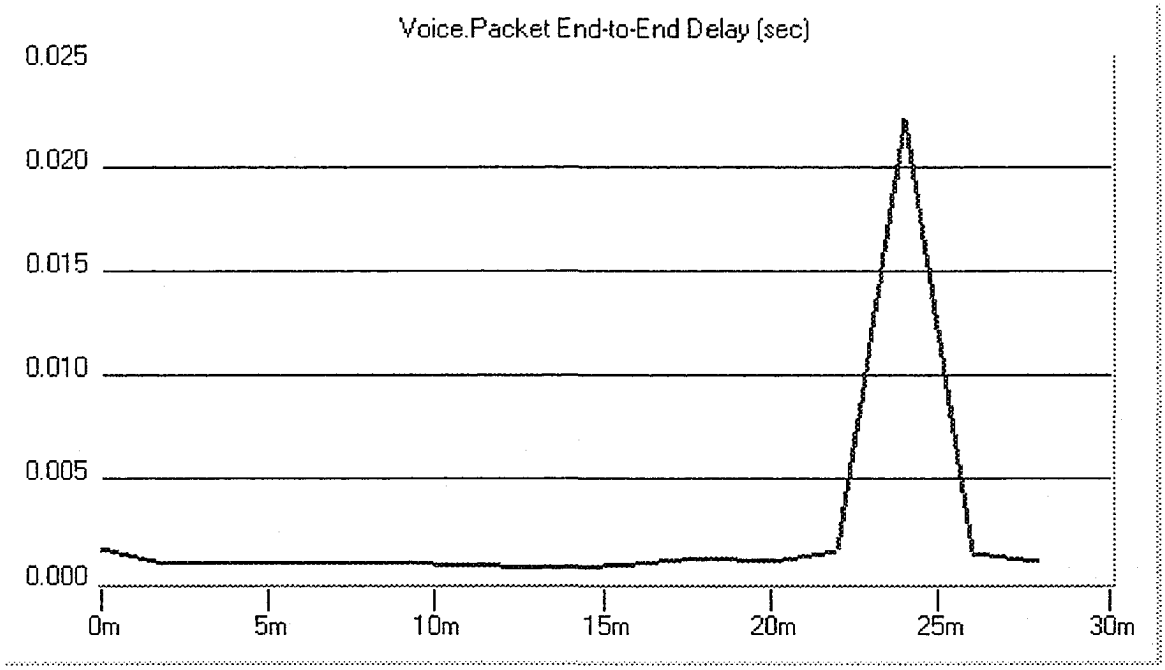
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Packet delay variation:

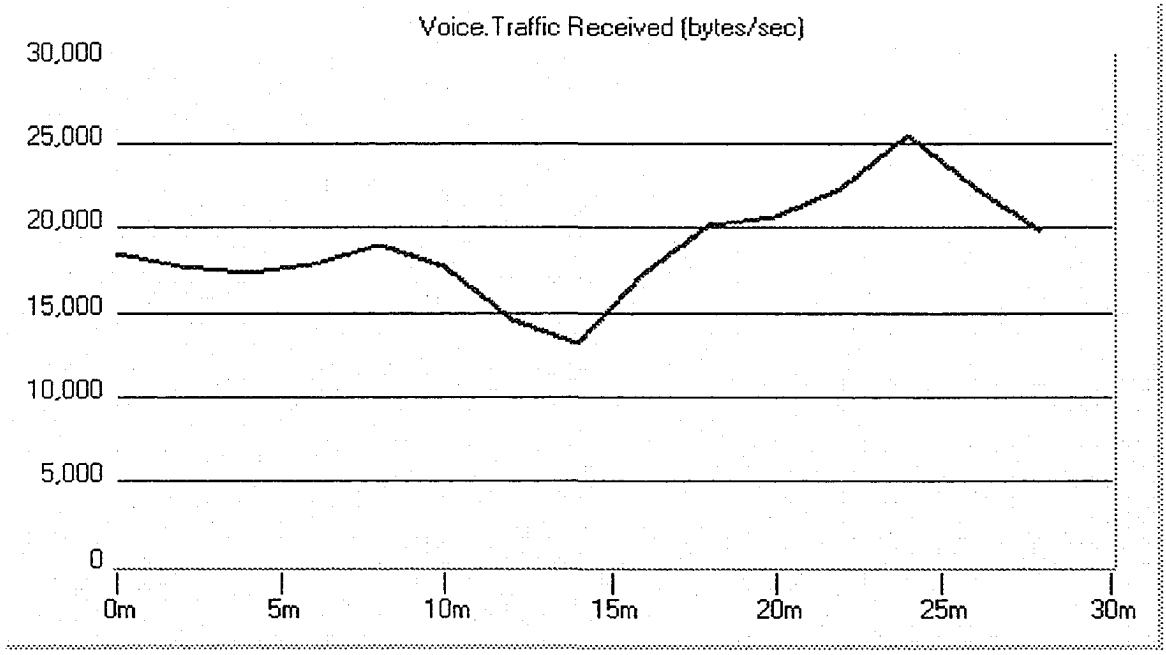


- Graph representation of Voice Packet End-to-End Delay:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

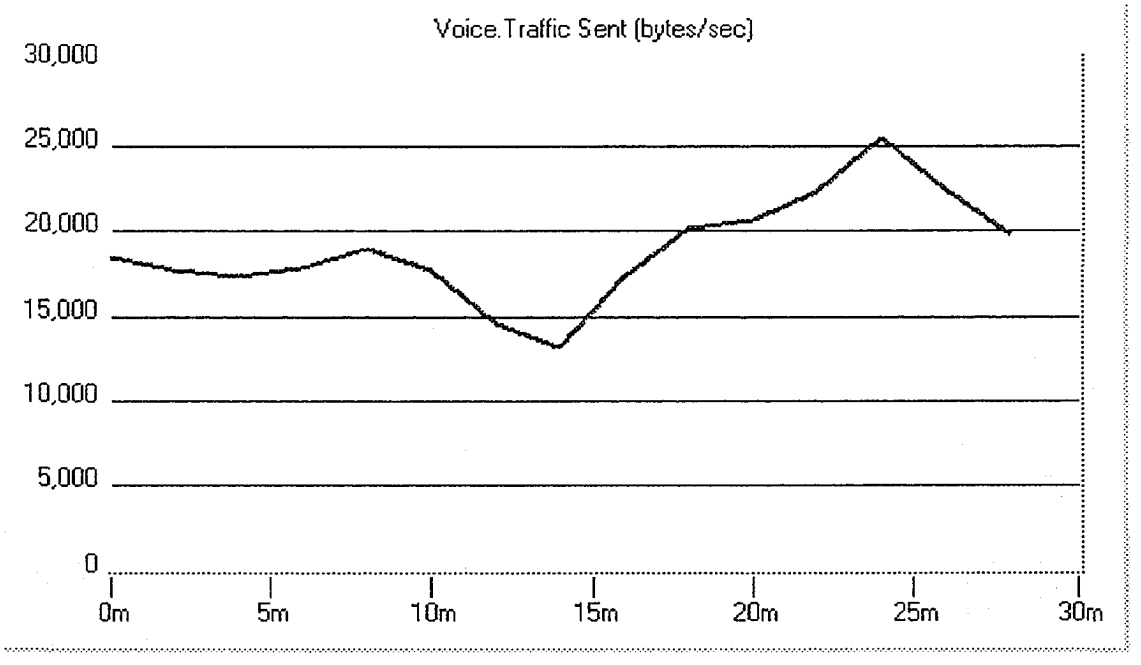


- Graph representation of Voice Traffic Received:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Traffic Sent:



- The relevant simulated parameters of the WLAN are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	23	276	0
Wireless LAN Delay (sec)	0.0019	0.0217	0.0002
Wireless LAN Load (bits/sec)	571,949	770,633	396,236
Wireless LAN Media Access Delay (sec)	0.0018	0.0207	0.0002
Wireless LAN Throughput (bits/sec)	571,927	770,364	396,236

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

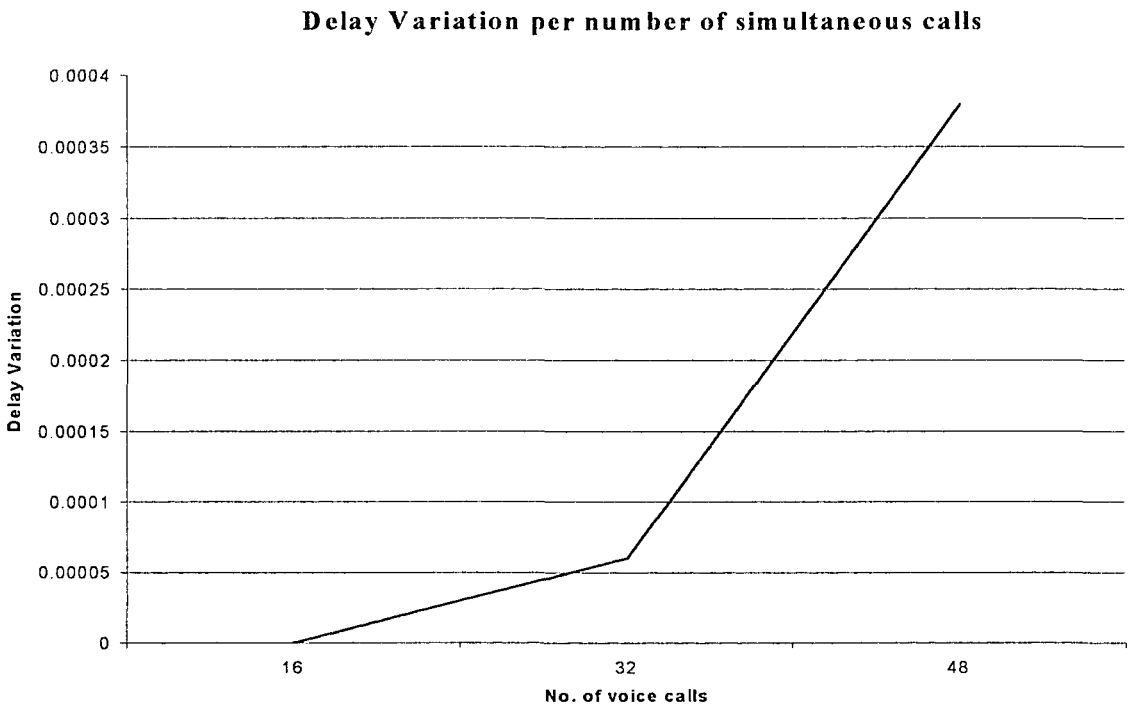
- Comparison table for number of simultaneous calls:

	16 wkst.	32wkst.	48wkst.
Statistic	Average	Average	Average
Voice Packet Delay Variation (sec)	0.00000011	0.00006	0.00038
Voice Packet End-to-End Delay (sec)	0.000591	0.00083	0.0024
Voice Traffic Received (bytes/sec)	10,730	17,596	18,811
Voice Traffic Sent (bytes/sec)	10,731	17,599	18,811

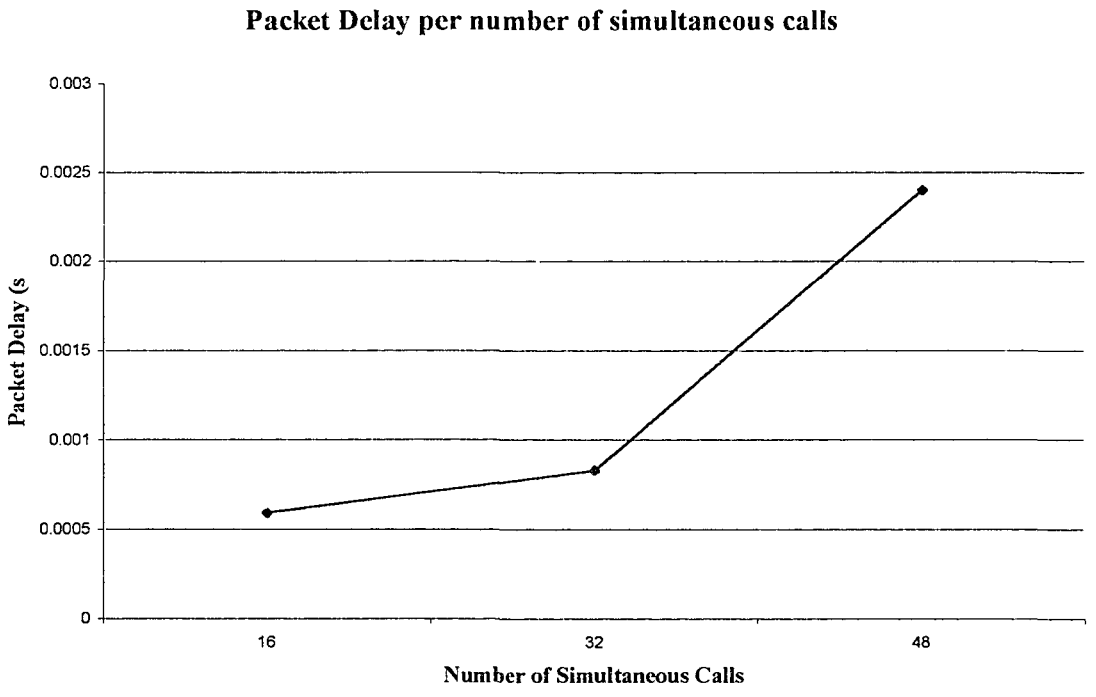
In a graphical representation of the table above, we can get a better understanding of the relationship between the degradation of the delay and delay variation as the number of simultaneous calls increases:

- Packet delay variation per number of simultaneous calls:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- Delay per number of simultaneous calls:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Conclusions:

- We can clearly notice a degradation of the voice performance (delay and delay variation) with the increase of the number of simultaneous voice calls

- We can notice that at a 50% increase of the number of simultaneous calls (from 32 to 48), the end-to end delay increases 4 times and the delay variation 6 times.

- Due to limited computer power, we were not able to successfully run significant simulations for higher than 48 wireless workstations, therefore we are not able to determine the number of simultaneous calls from where the voice quality parameters become unacceptable.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

II. Infrastructure Networks

As already described in detail, infrastructure networks use an Access Point for communication between wireless stations.

5th Scenario

In this scenario we study the impact of using an Access Point, in an infrastructure WLAN, on the performance of voice transmission.

With this goal in mind, we set up an infrastructure WLAN, with the following characteristics:

- 802.11b at 11Mbps
- Infrastructure network with one Access Point
- Physical Layer protocol is Direct Sequencing
- Calling scheme is as follows: workstation-1 calls workstation-2, workstation 3 calls workstation-4, ..., etc., bidirectional calls
- Voice codec used is G.729

We study the impact on the performance of voice transmission of the number of wireless workstations, in a generic infrastructure WLAN.

We simulate a infrastructure WLAN with one AP and 4, 6 and 8 workstations, and compare the voice performance statistics.

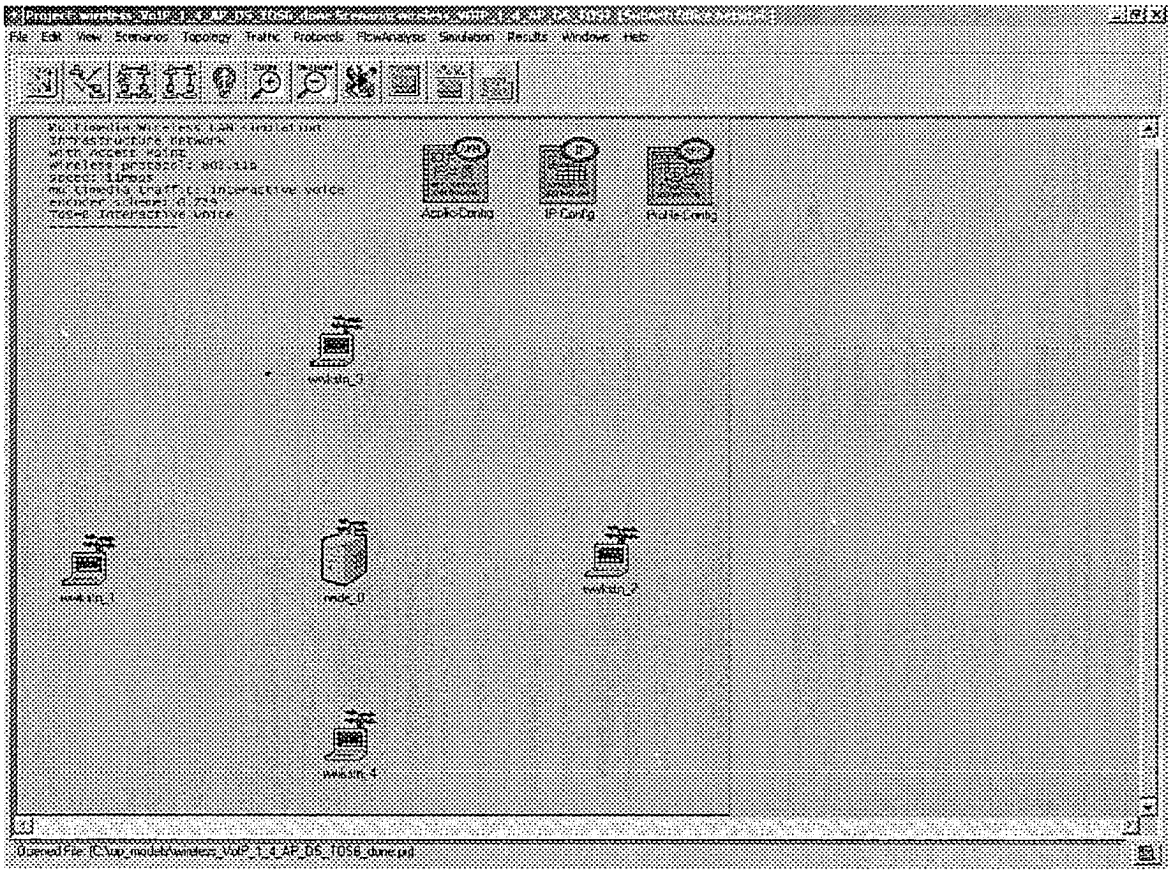
For this scenario a new network element, an Access Point was used. As Access Point, we used an already modeled wireless router with Access Point functionality enabled.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Because of the increased number of computations, as the number of workstations increases, we reduced simulation time to as low as 5 minutes and the number of collected statistics from 100 to as low as 50.

Subscenario 5.1 Simulation with 1 access point and 4 wireless workstations:

- The general architecture of this network is:



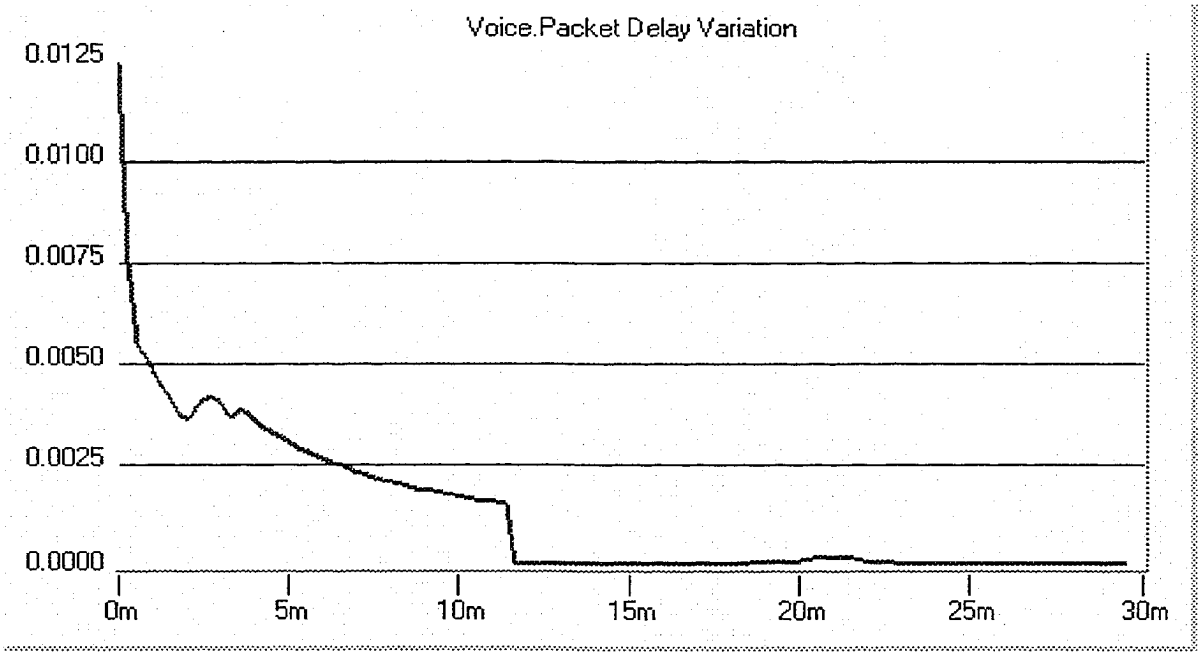
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

We simulated 30 minutes of voice traffic on the network.

- The simulated voice quality parameters are:

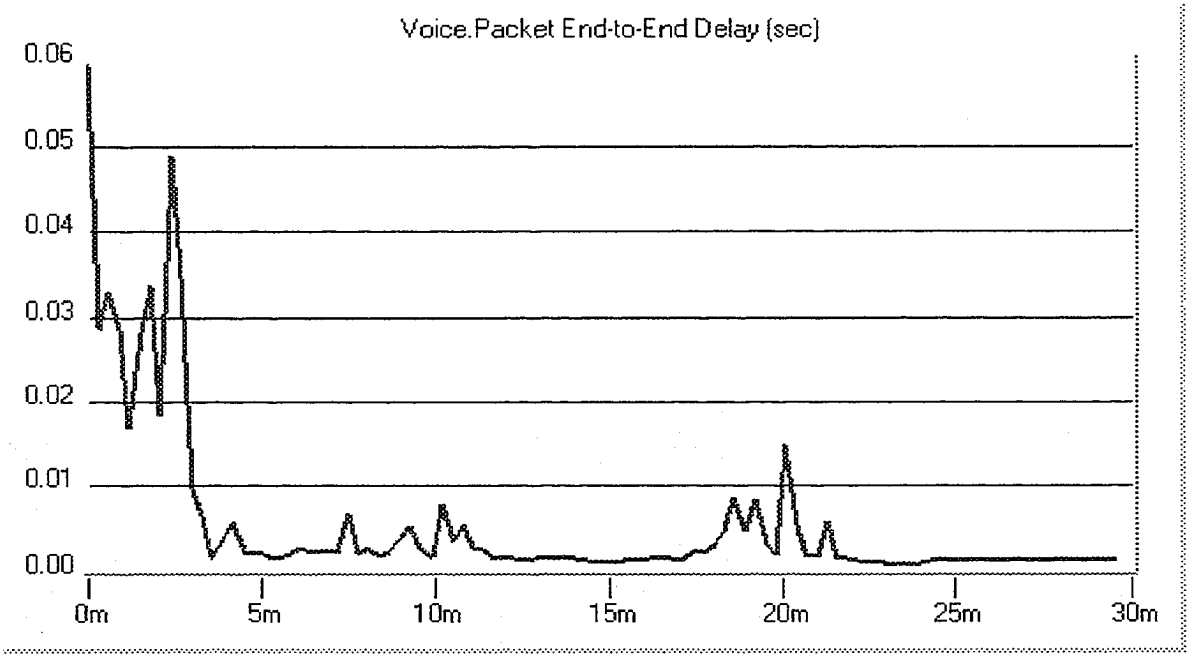
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.0013	0.0124	0.0000
Voice Packet End-to-End Delay (sec)	0.0053	0.0593	0.0004
Voice Traffic Received (bytes/sec)	2,110	3,919	374
Voice Traffic Sent (bytes/sec)	2,110	3,919	378

- Graph representation of Voice Packet Delay Variation:

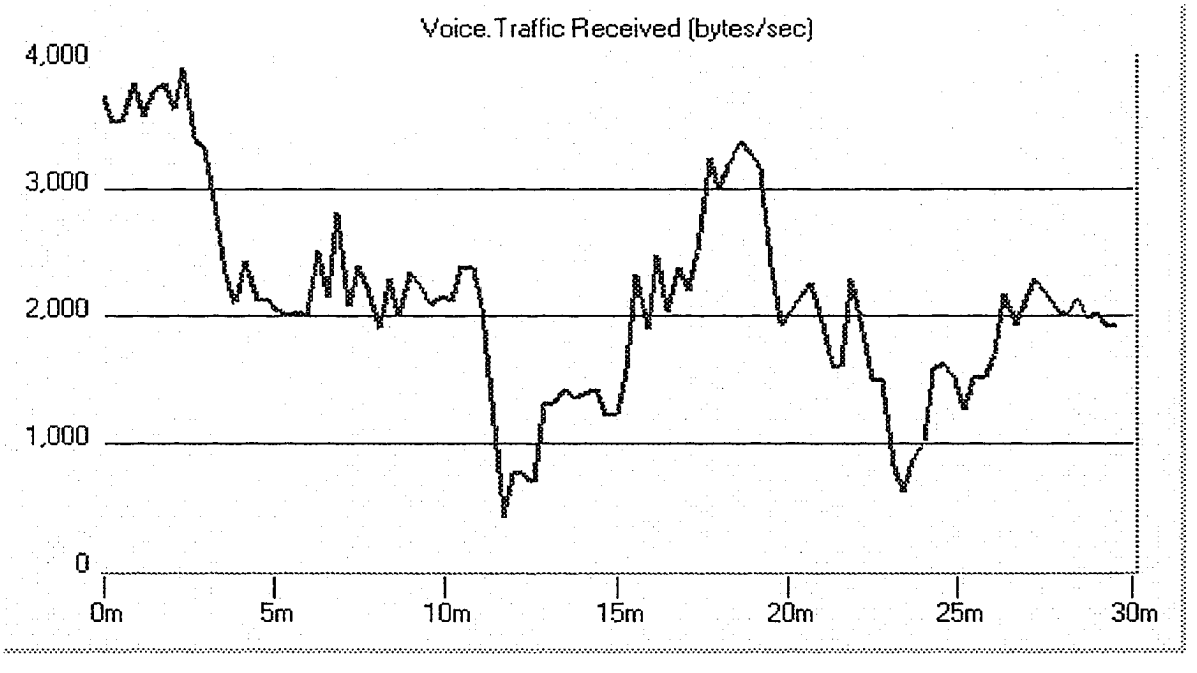


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Packet Delay:

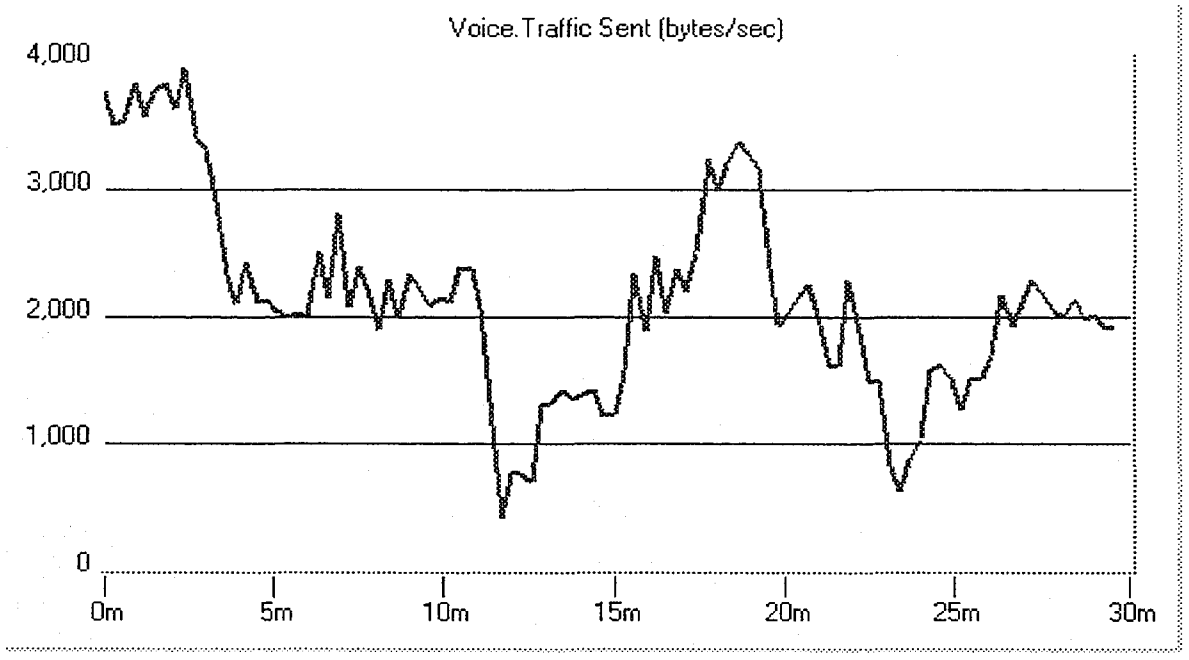


- Graph representation of Voice Traffic Received:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Traffic Sent:



- WLAN Traffic parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	0	0	0
Wireless LAN Delay (sec)	0.0049	0.0589	0.0000
Wireless LAN Load (bits/sec)	244,511	449,707	38,969
Wireless LAN Media Access Delay (sec)	0.0026	0.0319	0.0000
Wireless LAN Throughput (bits/sec)	135,046	250,809	23,993

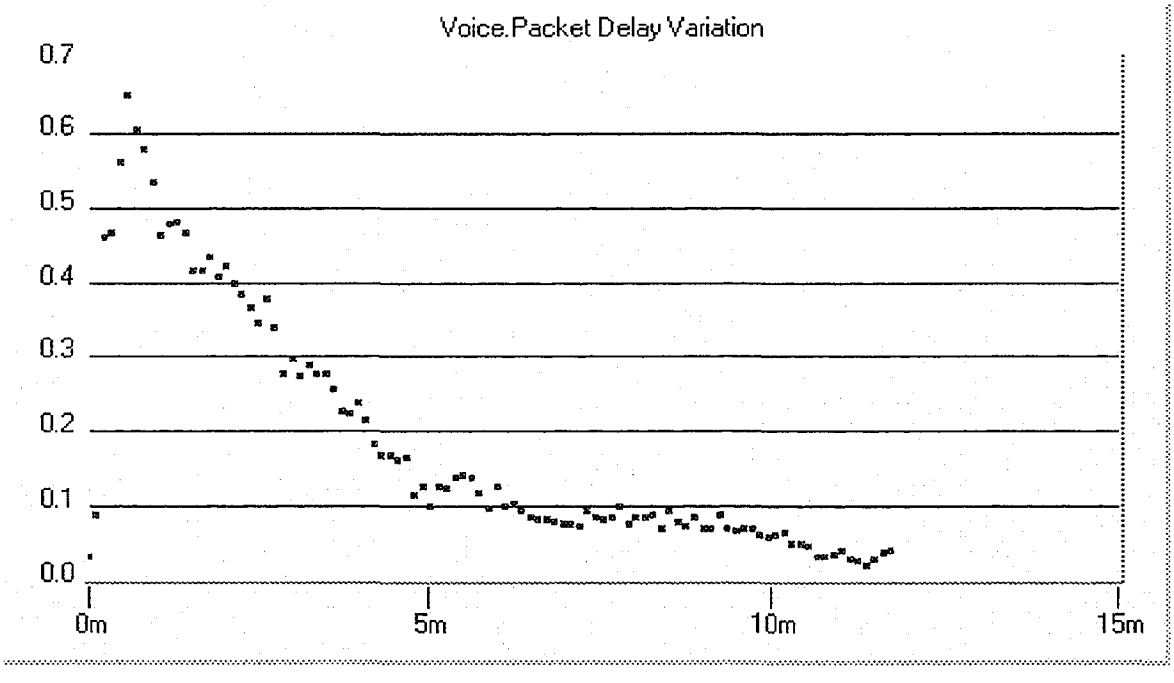
Subscenario 5.2: Simulation with 1 access point and 6 wireless fixed workstations:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Voice quality parameters are:

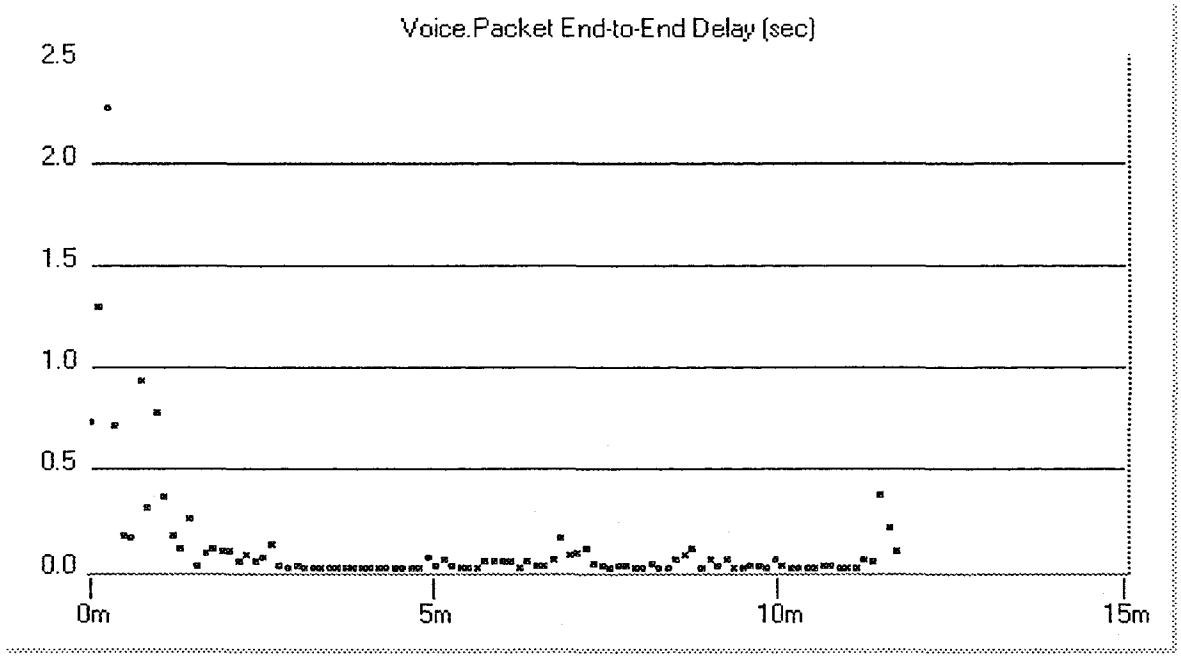
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.183	0.652	0.015
Voice Packet End-to-End Delay (sec)	0.11	2.26	0.00
Voice Traffic Received (bytes/sec)	3,528	5,340	1,936
Voice Traffic Received (packets/sec)	882	1,335	484
Voice Traffic Sent (bytes/sec)	3,530	5,925	1,935
Voice Traffic Sent (packets/sec)	882	1,481	484

- Graph representation of Voice Packet delay variation:

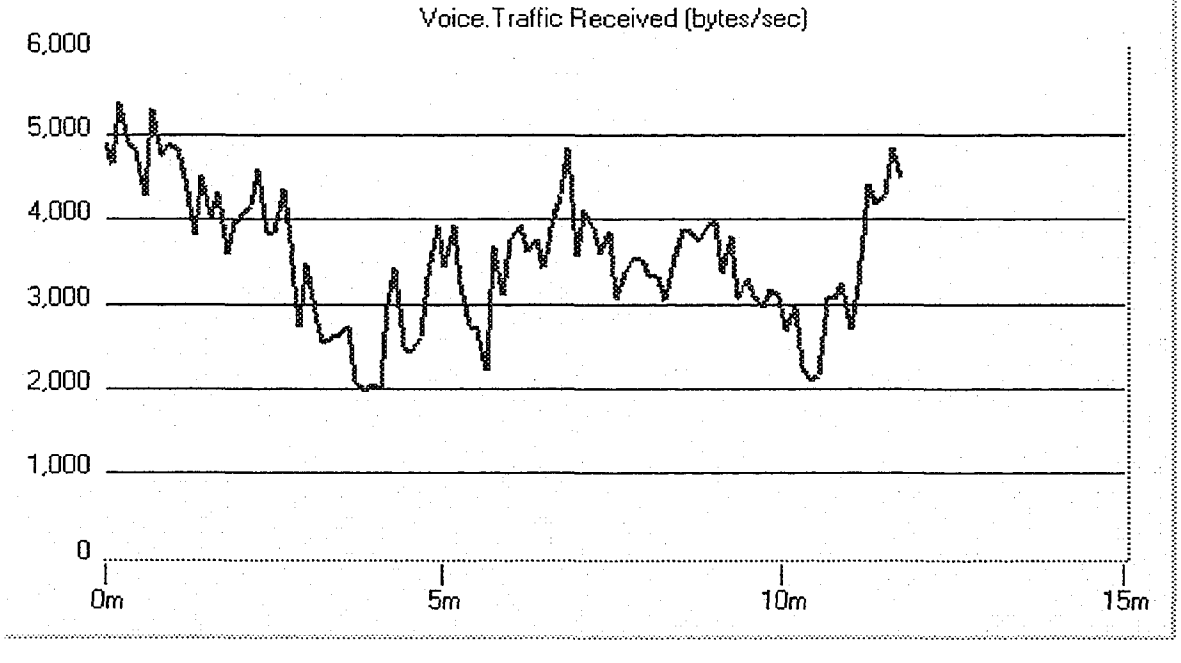


- Graph representation of Voice Packet End-to-End Delay:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

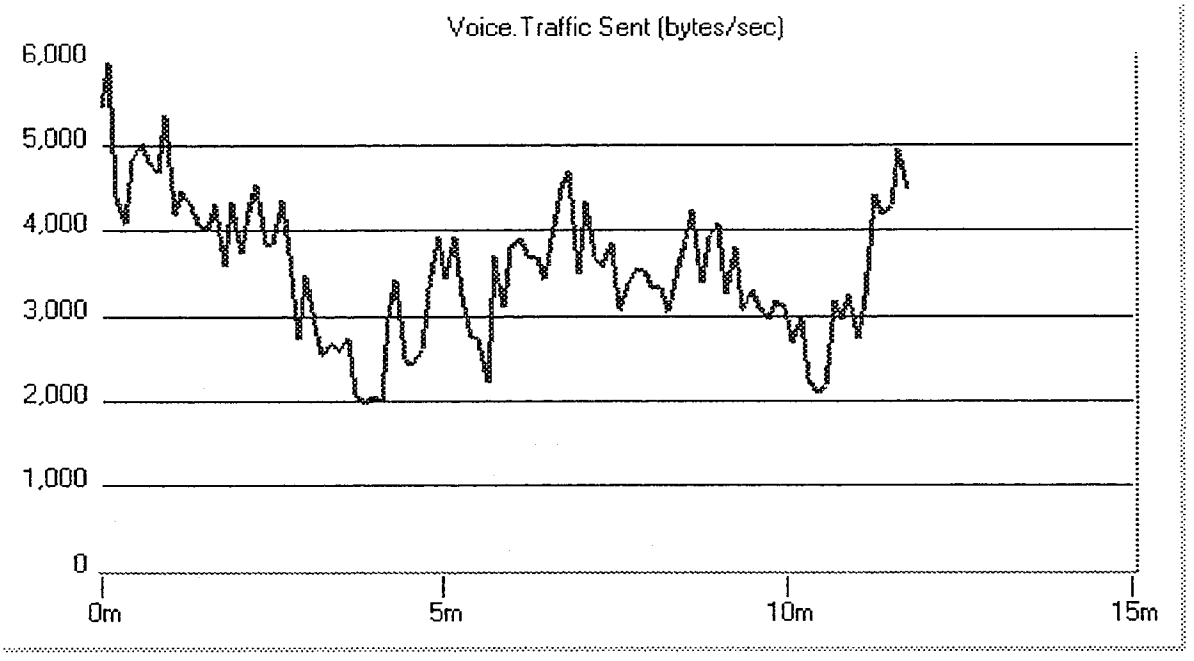


- Graph representation of Voice Traffic Received:



- Graph representation of Voice Traffic Sent

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Subscenario 5.3 With 1 Access Point and 8 wireless fixed workstations:

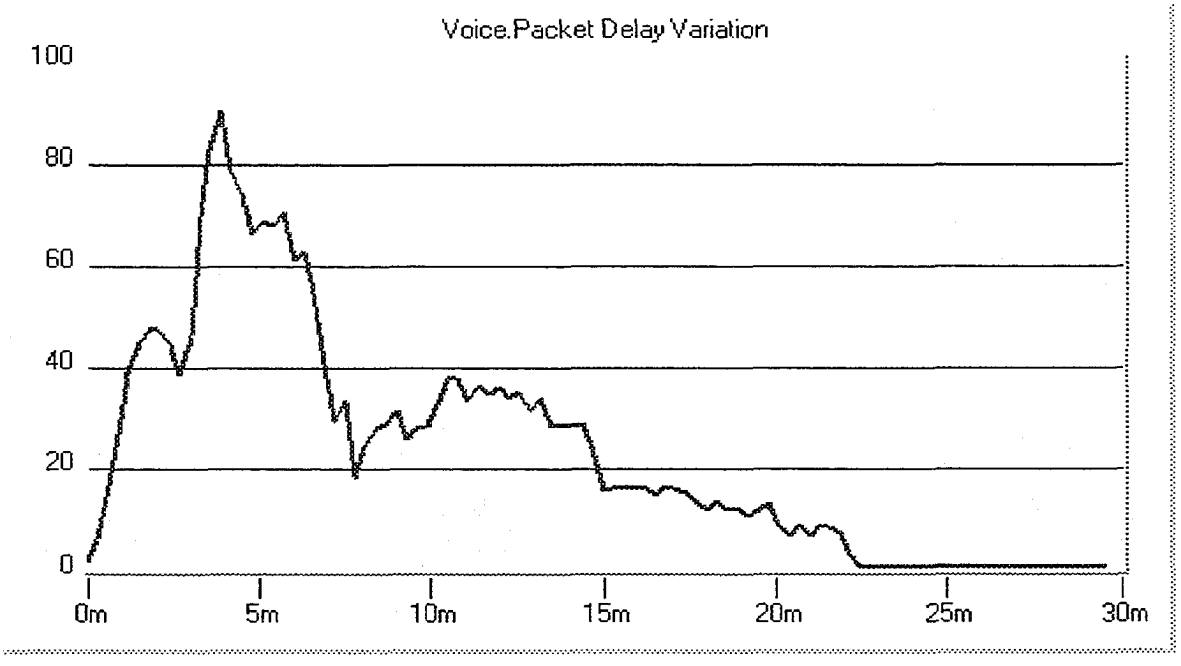
- Always we can clearly notice a degradation of the voice performance (delay and delay variation) with the increase of the number of workstations.

- Voice quality parameters are:

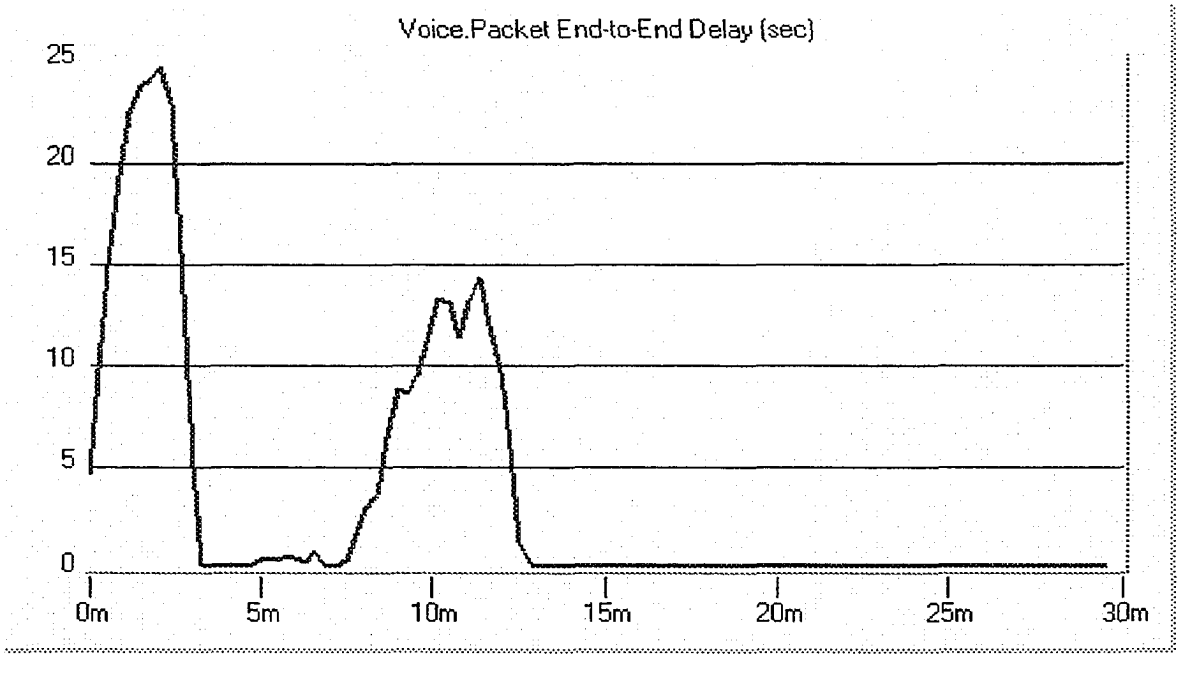
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	23.3	90.1	0.0
Voice Packet End-to-End Delay (sec)	3.4	24.6	0.0
Voice Traffic Received (bytes/sec)	3,235	5,413	670
Voice Traffic Sent (bytes/sec)	3,235	6,199	670

- Graph representation of Voice Packet delay variation:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

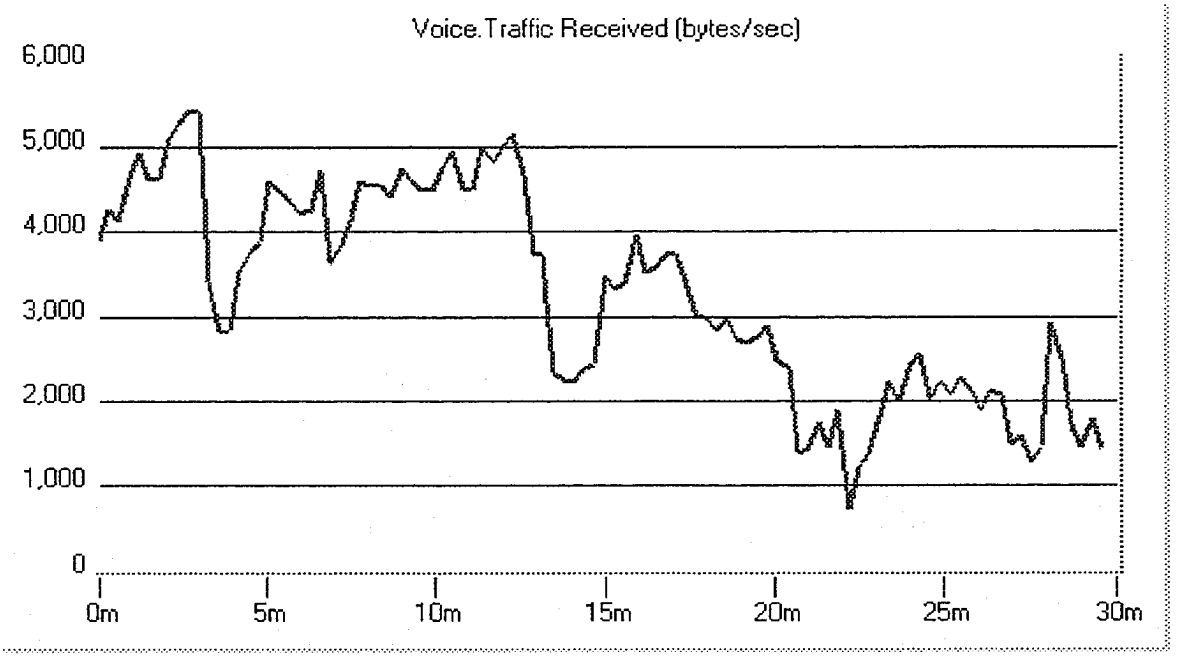


- Graph representation of Voice Packet End-to-End Delay:

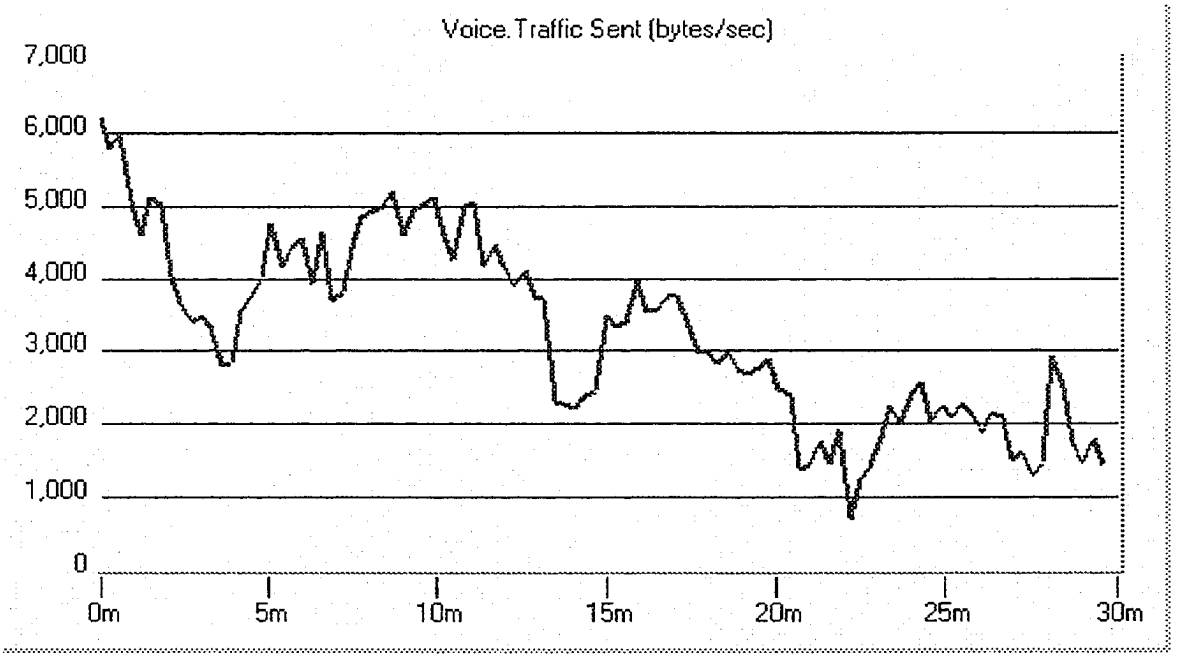


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Traffic Received:



- Graph representation of Voice Traffic Sent:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Simulated WLAN parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	0	0	0
Wireless LAN Delay (sec)	0.11	2.26	0.00
Wireless LAN Load (bits/sec)	415,168	704,818	201,636
Wireless LAN Media Access Delay (sec)	0.06	1.31	0.00
Wireless LAN Throughput (bits/sec)	225,817	341,760	123,840

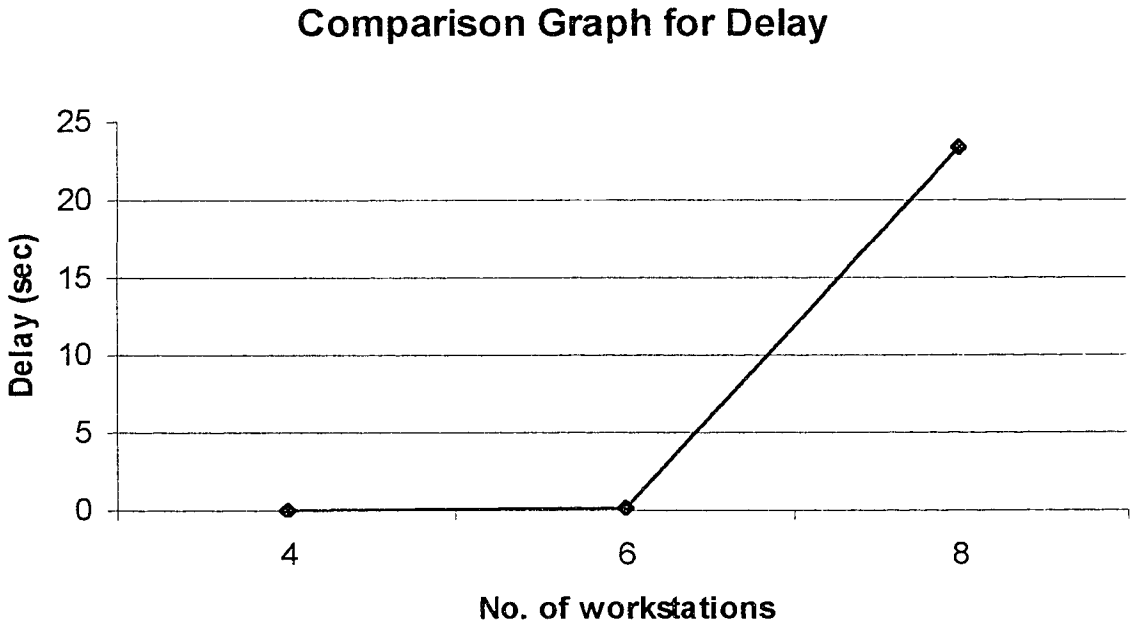
- Comparison table for infrastructure network by number of simultaneous calls:

	1AP+4wkst.	1AP+6wkst.	1AP+8wkst.
Statistic	Average	Average	Average
Voice Packet Delay Variation	0.0013	0.1830	23.3000
Voice Packet End-to-End Delay	0.0053	0.1100	3.400
Voice Traffic Received (bytes/sec)	2,110	3,528	3,235
Voice Traffic Sent (bytes/sec)	2,110	3,530	3,235

-Comparison Graph for End-to-end Delay

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Comparison graph for Delay Variation:



Conclusions for Scenario 5:

- We can notice that in an infrastructure wireless network, using an Access Point for peer to peer communication, the voice quality parameters become prohibitively high at even a small number of simultaneous calls.

- The degradation happens even at a small increase of the number of simultaneous calls.

- If, for four simultaneous calls, delay and delay variation is in acceptable range, at six simultaneous calls it reaches the acceptable limits and in the case of eight simultaneous calls it is well above the acceptable limits (limits presented at the beginning of this chapter).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- So, we can conclude, that in an infrastructure WLAN, only up to six simultaneous calls are qualitatively acceptable.

III. Internetworked Wireless LANs

In this category of simulations we consider two or several wireless local area networks interconnected with wireline connections or internetworked with wireline networks. We take into consideration 3 differnet type of interconnections:

- Fast Ethernet at a speed of 100Mbps, with a maximum of 300 ft connection distance (used usually in a campus internetwork)

- Wide Area Network connection with bidirectional DS1 line, working at a speed of 1.54 Mbps, usually used to interconnect remote campuses,

- Wide Area Network connection through an Internet cloud, which gives the possibility of world-wide interconnection.

- RIP v1 is modeled and used as routing protocol, for all internetworked WLANs

6th Scenario

In this scenario we study the impact of the number of router or the number wireless workstations, on the performance of voice transmission between two WLANs, interconnected with wireline network.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

With this goal in mind, we set up two identical Wireless LANs interconnected with routers, having the following characteristics:

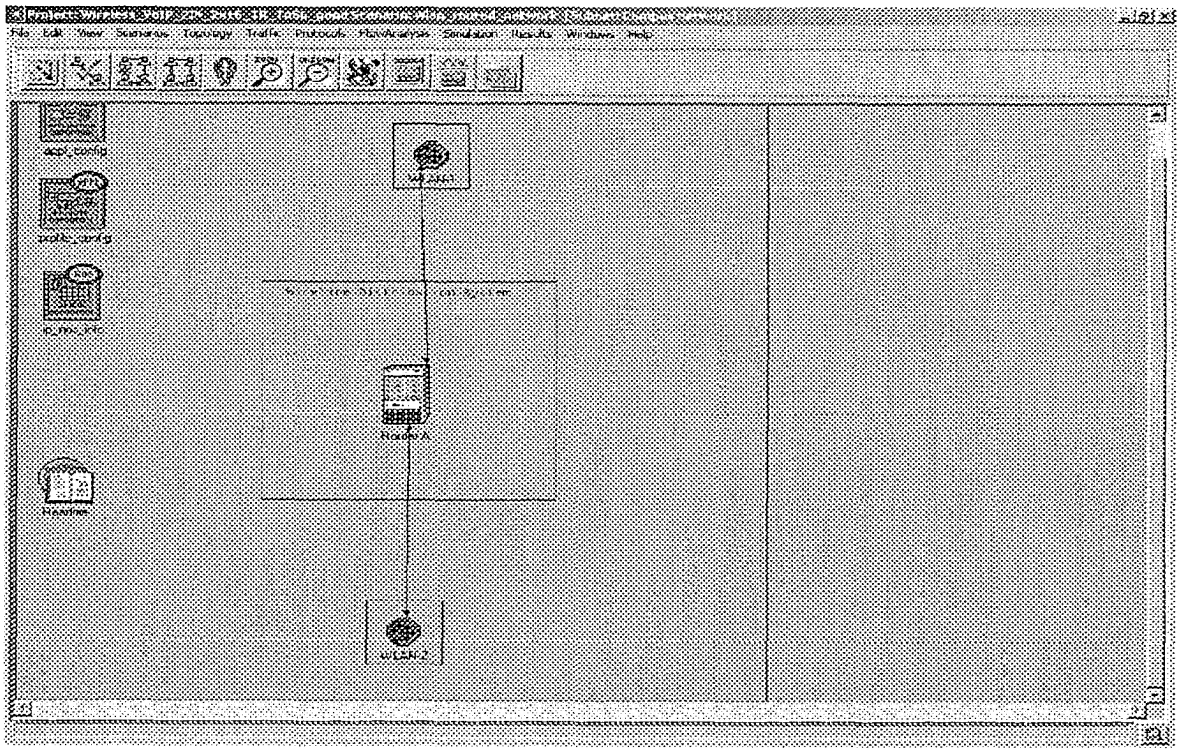
- 802.11b at 11Mbps
- Infrastructure network with an Access Point
- Physical Layer protocol is Direct Sequencing
- Voice is routed between the two WLANS through wireline routers
- 16 simultaneous bi-directional packetized voice calls
- Calling scheme is as follows: workstation-1 calls workstation-2, workstation 3 calls workstation-4, ..., etc., with bidirectional traffic
- Voice codec used is G.729
- Wireline interconnection using 100Mbps Ethernet connection between the access points (AP) and the router.

We study the impact on the performance of voice transmission between the two WLANS, when we increase the number of routers in the internetworked backbone.

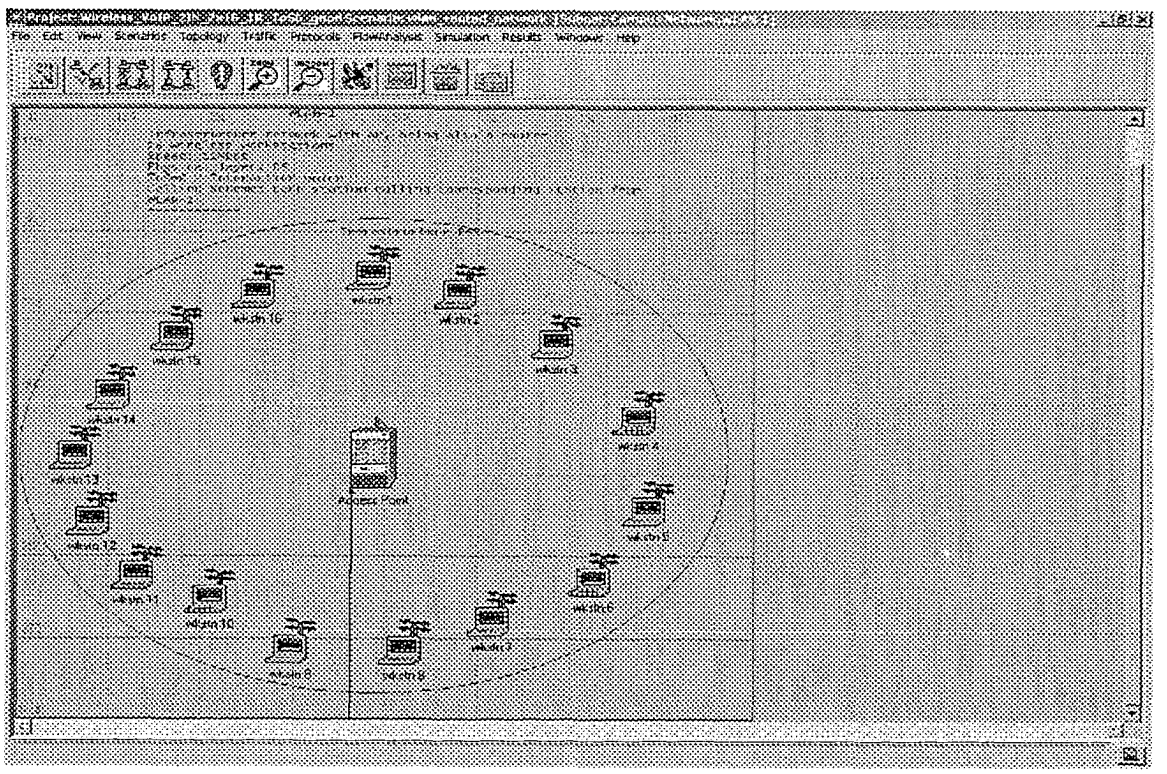
Sub-scenario 6.1 : One wireline router connecting two wireless LANs

- General Architecture of the interconnected WLANs:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- General Architecture of one WLAN:



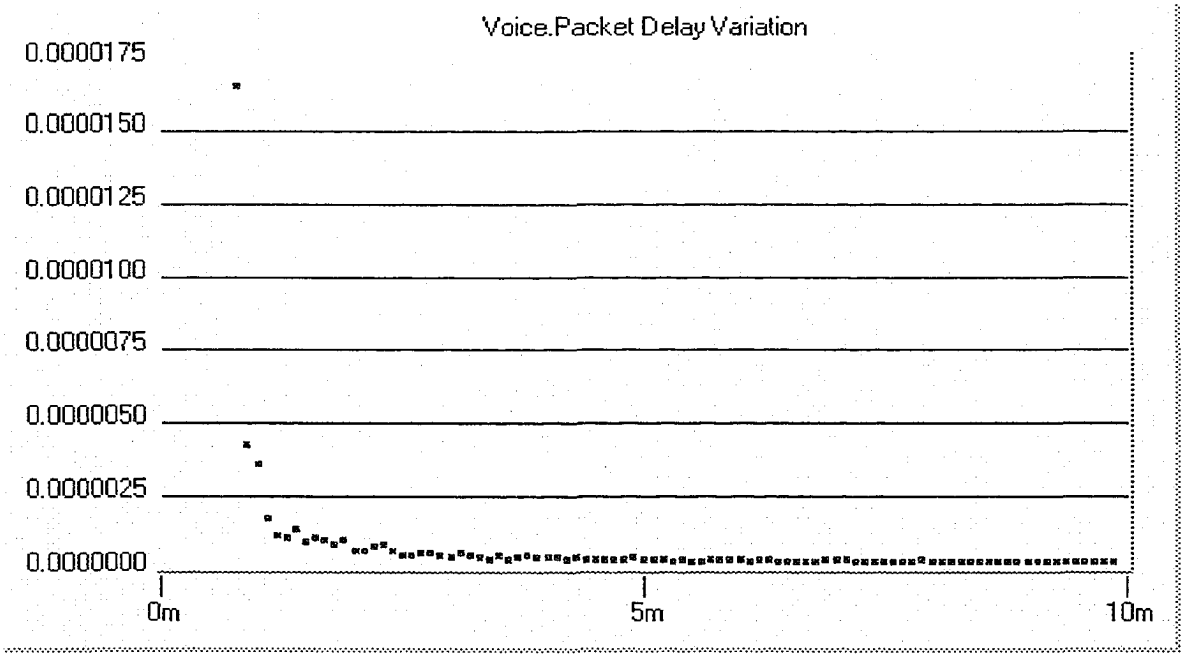
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Simulation Results are:

- Voice quality parameters are:

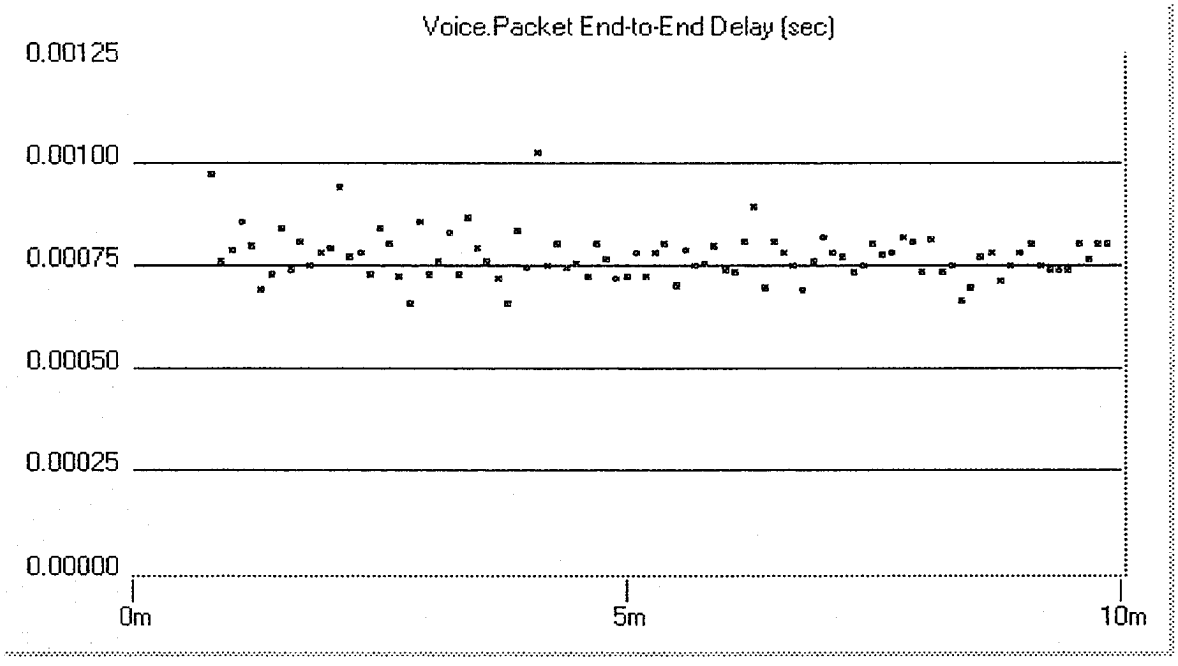
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.0000006	0.0000165	0.0000002
Voice Packet End-to-End Delay (sec)	0.00077	0.00102	0.00065
Voice Traffic Received (bytes/sec)	2,700	4,007	0
Voice Traffic Sent (bytes/sec)	2,700	4,006	0

- Graph representation of Voice Packet delay variation:



- Graph representation of Voice Packet End-to-End Delay:

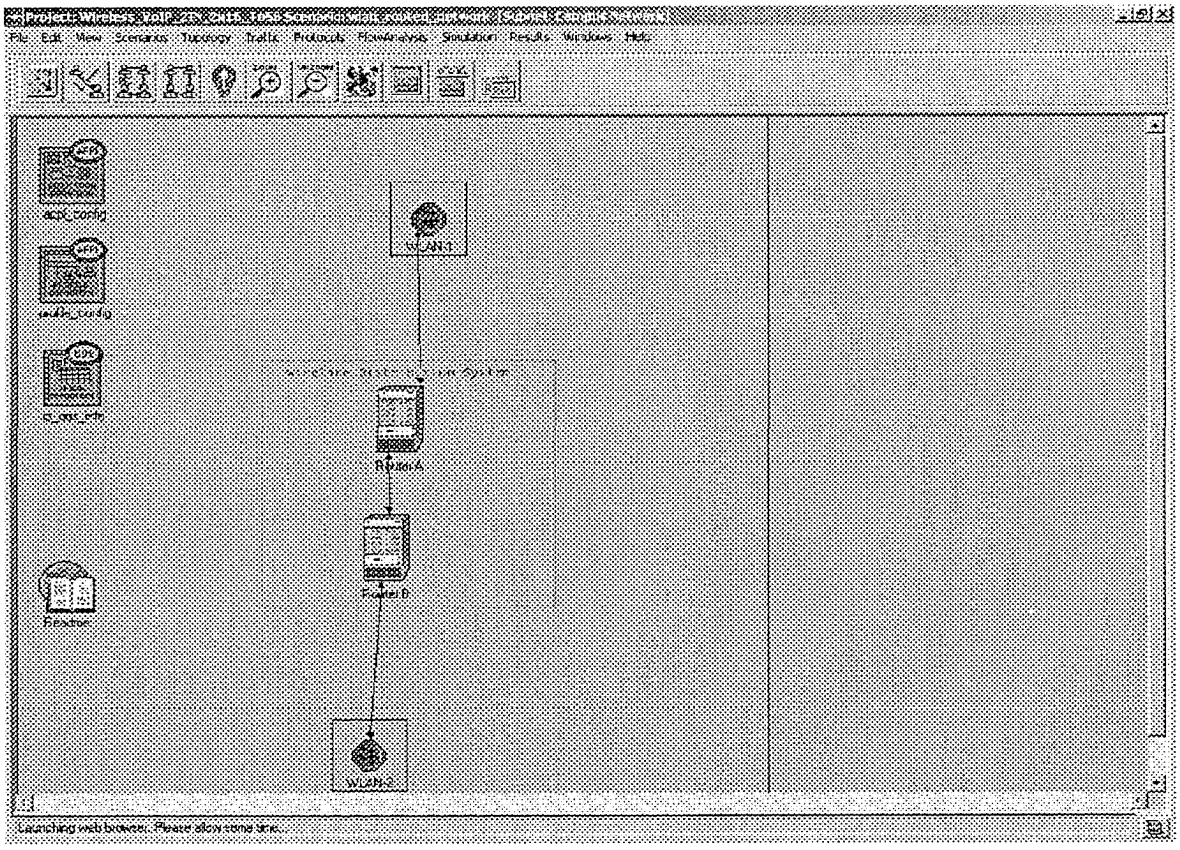
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Sub-scenario 6.2 In this sub-scenario we consider **two routers**, interconnecting two WLANs, with 16 wireless workstations each. All other parameters of the network remain unchanged.

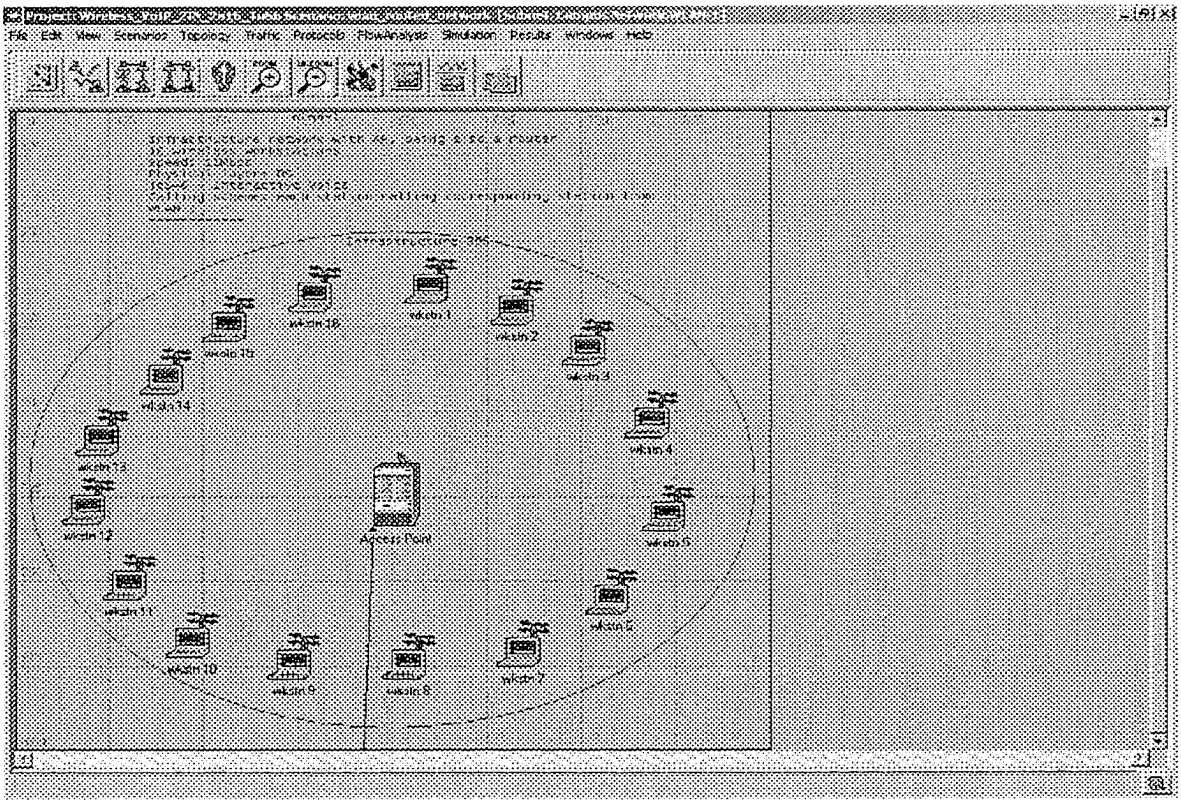
- The General Architecture of the network is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- Each WLAN architecture is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

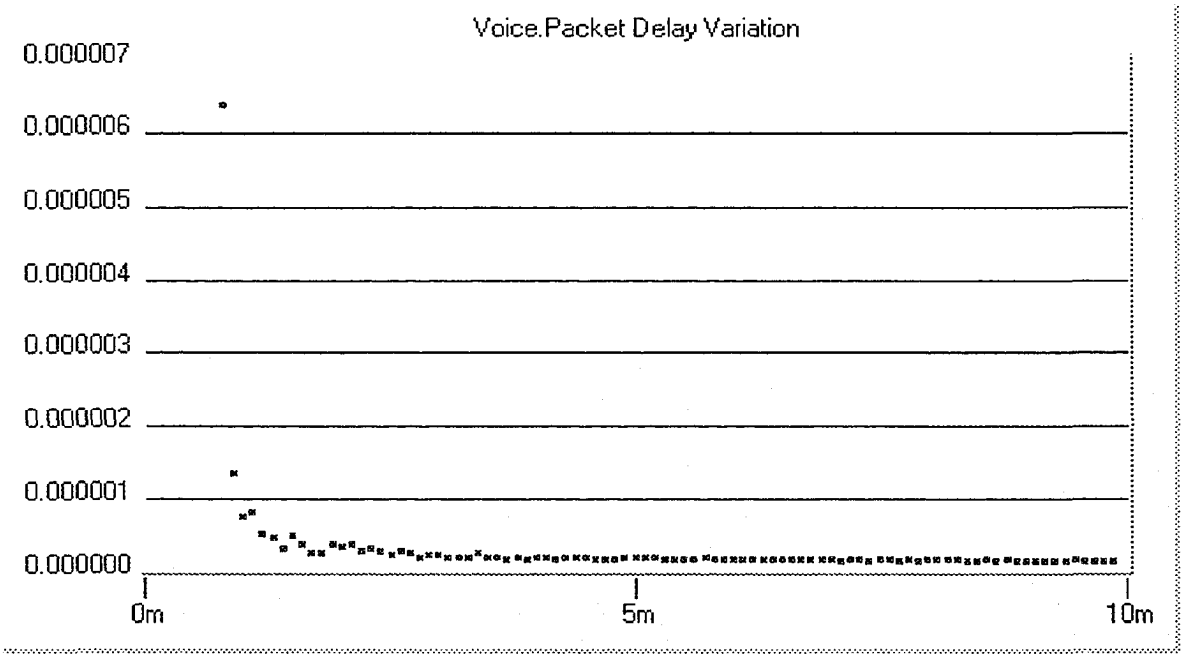


- Voice quality parameters are:

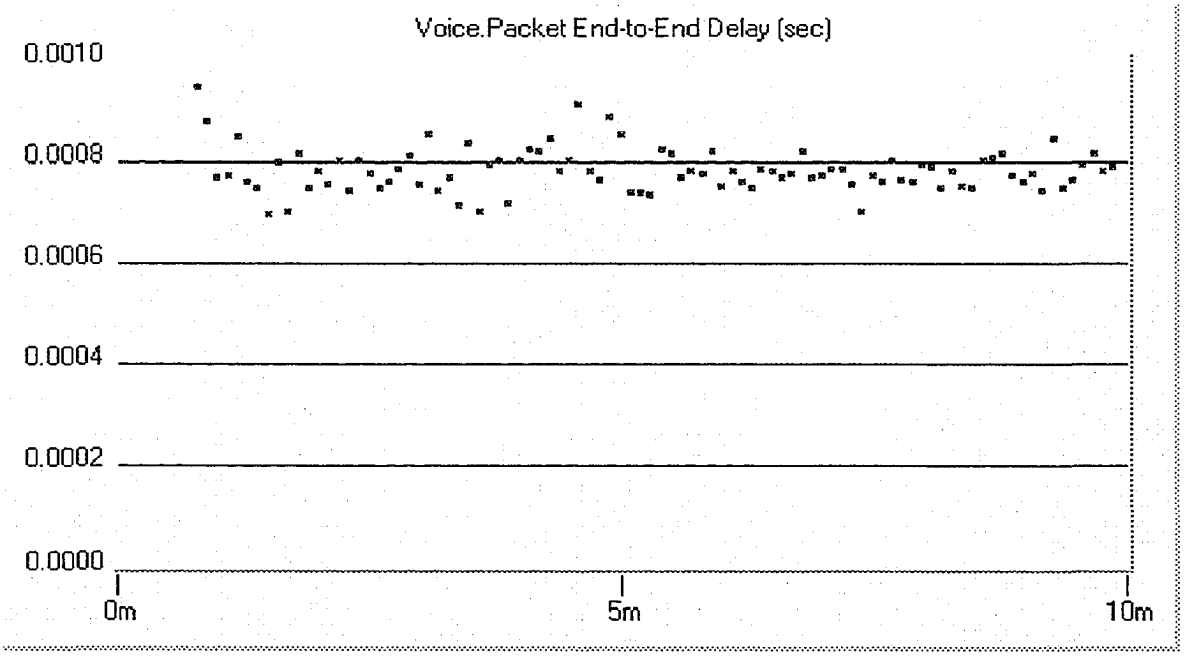
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.00000026	0.00000637	0.00000011
Voice Packet End-to-End Delay (sec)	0.000779	0.000941	0.000692
Voice Traffic Received (bytes/sec)	2,601	3,929	0
Voice Traffic Sent (bytes/sec)	2,602	3,928	0

- Graph representation of Voice Packet delay variation:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

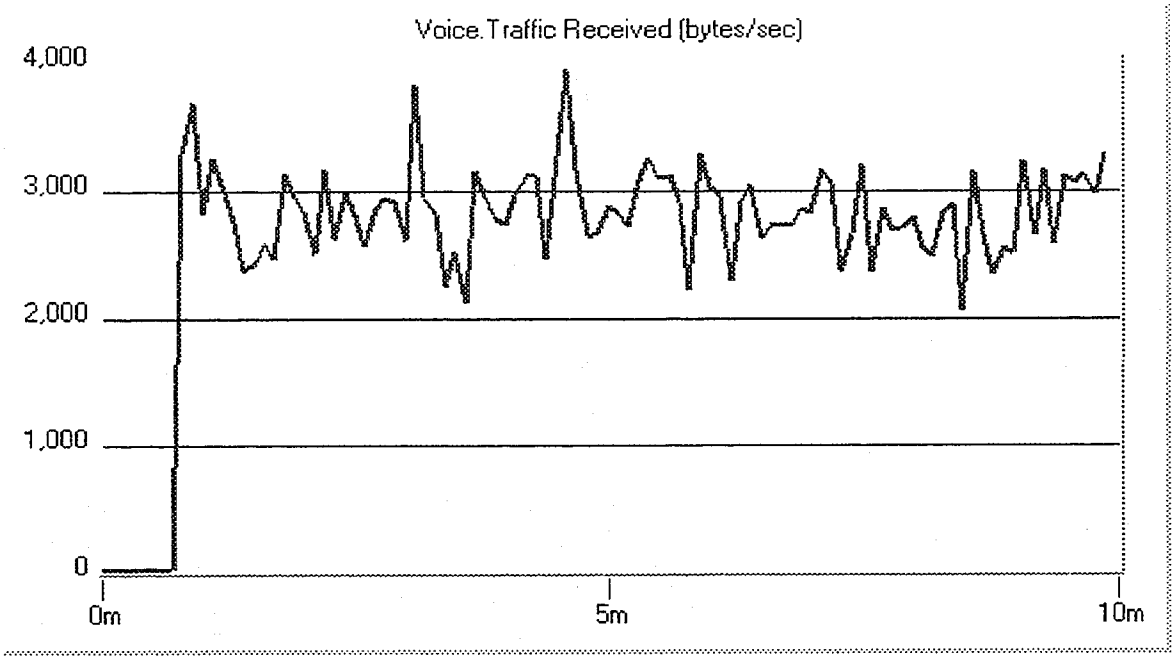


- Graph representation of Voice Packet End-to-End Delay:

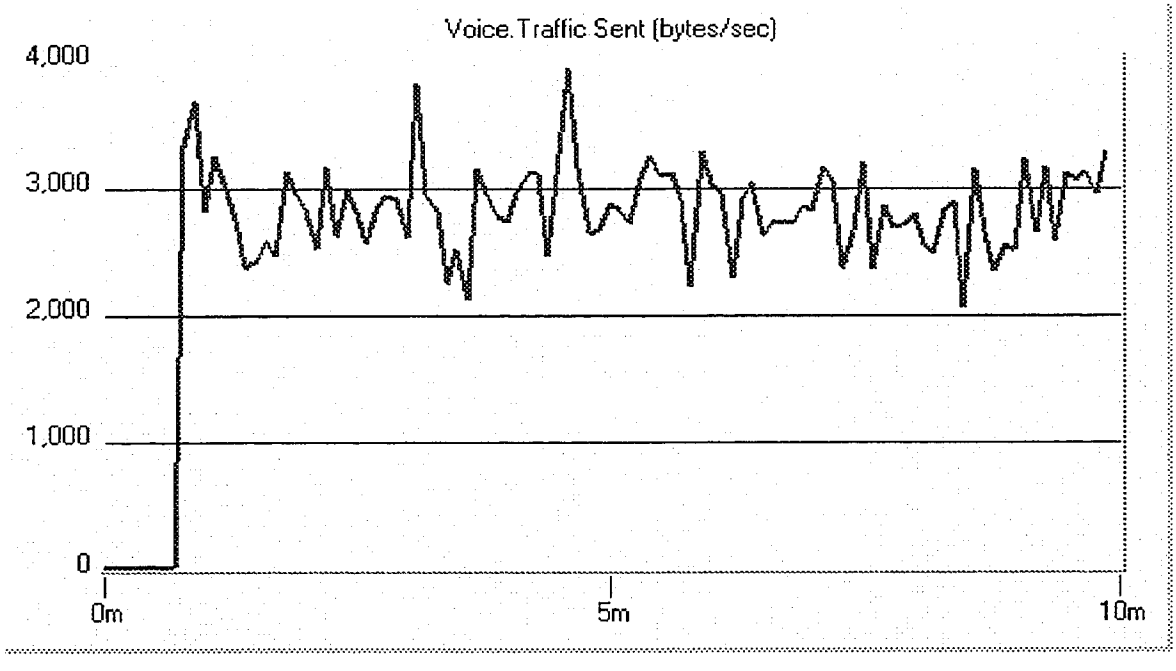


- Graph representation of Voice Traffic Received:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- Graph representation of Voice Traffic Sent:



- WLAN traffic parameters are:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	13	1,344	0
Wireless LAN Delay (sec)	0.000115	0.000209	0.000075
Wireless LAN Load (bits/sec)	333,116	502,827	0
Wireless LAN Media Access Delay (sec)	0.000066	0.000162	0.000000
Wireless LAN Throughput (bits/sec)	333,332	502,827	0

- Comparison Table for 1 vs 2 routers in internetworked WLANs:

	1 Router	2 Routers
Statistic	Average	Average
Voice Packet Delay Variation	0.0000006	0.00000026
Voice Packet End-to-End Delay (sec)	0.00077	0.000779
Voice Traffic Received (bytes/sec)	2,700	2,601
Voice Traffic Received (packets/sec)	2,700	6,605

Conclusions for adding a second router:

Comparing the results from subscenarios 6.1 and 6.2 we can conclude that by adding a second router to the simulated network has no significant impact on the voice quality transmission, in case the same number of 16 simultaneous calls are generated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Scenario 7:

In this scenario we study the impact on the quality of voice transmission over internetworked WLANs, in the case we split each WLAN in 2 equal sub-LANs.

Sub-scenario 7.1

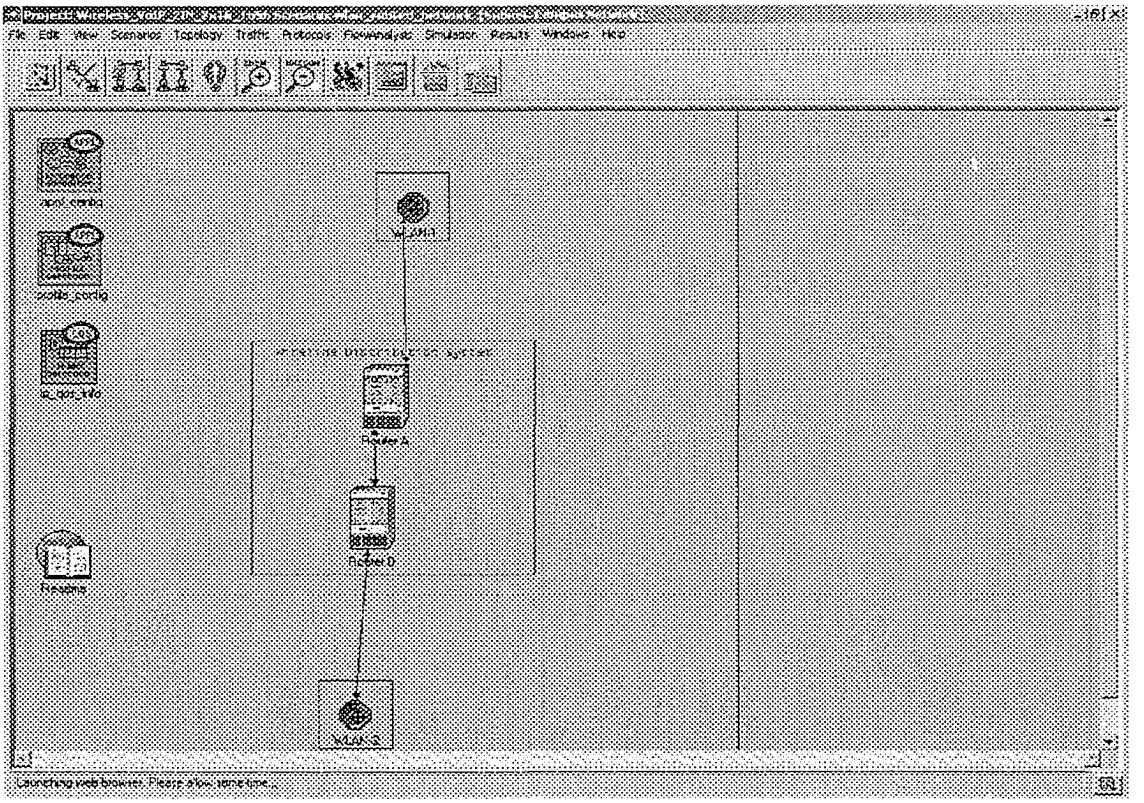
With this goal in mind, we set up two identical Wireless LANs interconnected with routers, having the following characteristics:

- 802.11b at 11Mbps
- Infrastructure network with an Access Point
- Physical Layer protocol is Direct Sequencing
- Voice is routed between the two WLANS through wireline routers
- Calling scheme is as follows: workstation-1 calls workstation-2, workstation 3 calls workstation-4, ..., etc., with bidirectional traffic
- Voice codec used is G.729
- Wireline interconnection using 100Mbps Ethernet connection between the access points (AP) and the router.

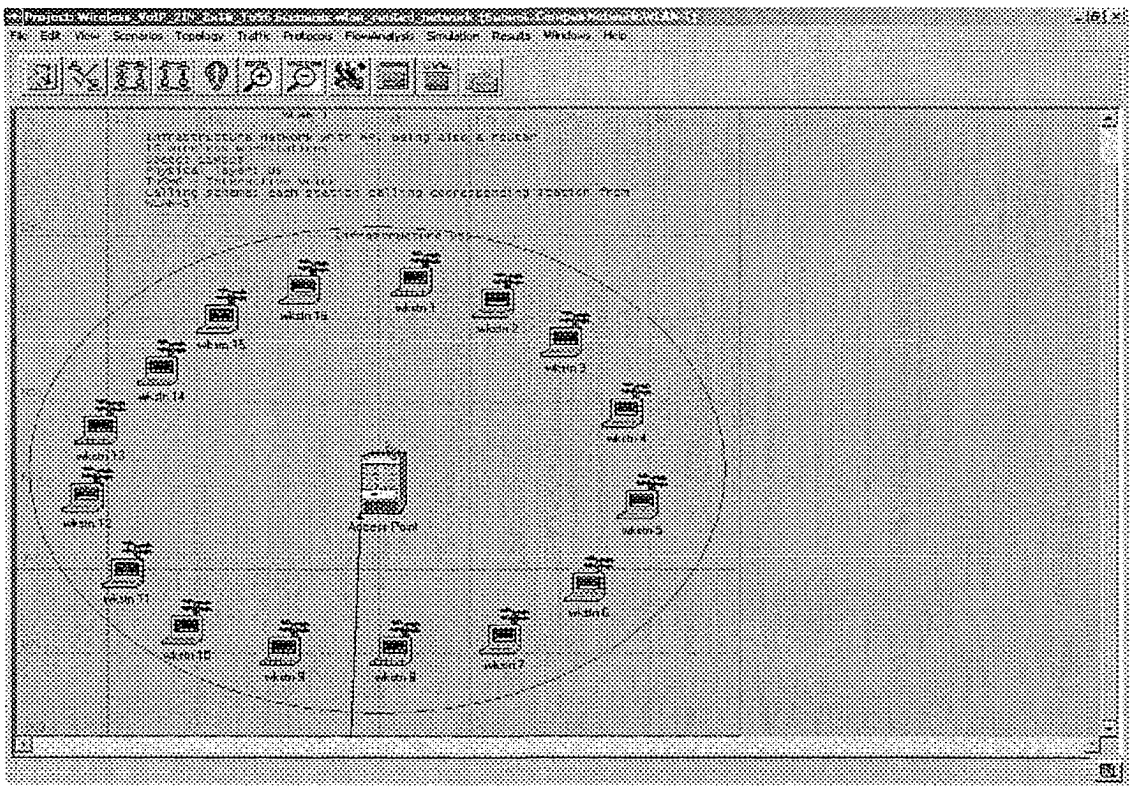
This subscenario is identical with Subscenario 6.2

- The general architecture of Subscenario 7.1 is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- The architecture of each of the 2 WLANs is:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

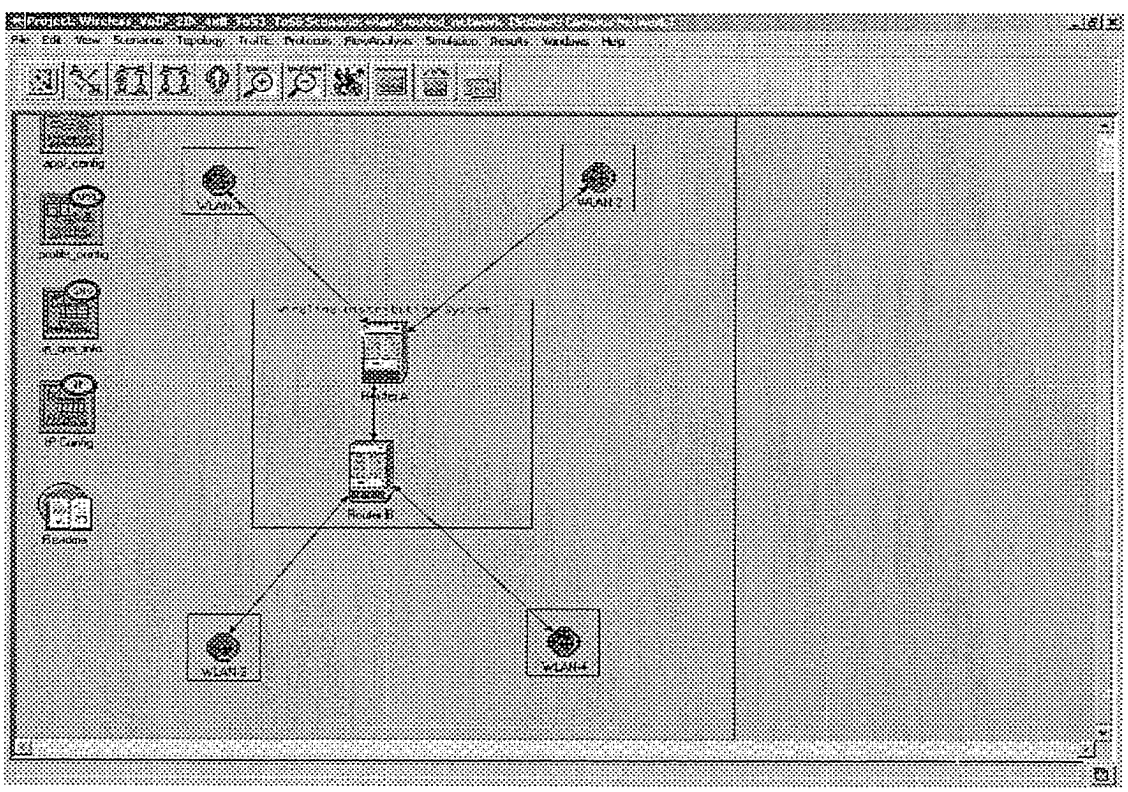
Sub-scenario 7.2

In this sub-scenario we keep the backbone the same from the previous subscenario 7.1 and split each WLAN (of 16 wireless workstations) in 2 identical sub-WLANs of 8 wireless workstations each.

We study the impact of splitting the WLANS on the quality of voice transmission.

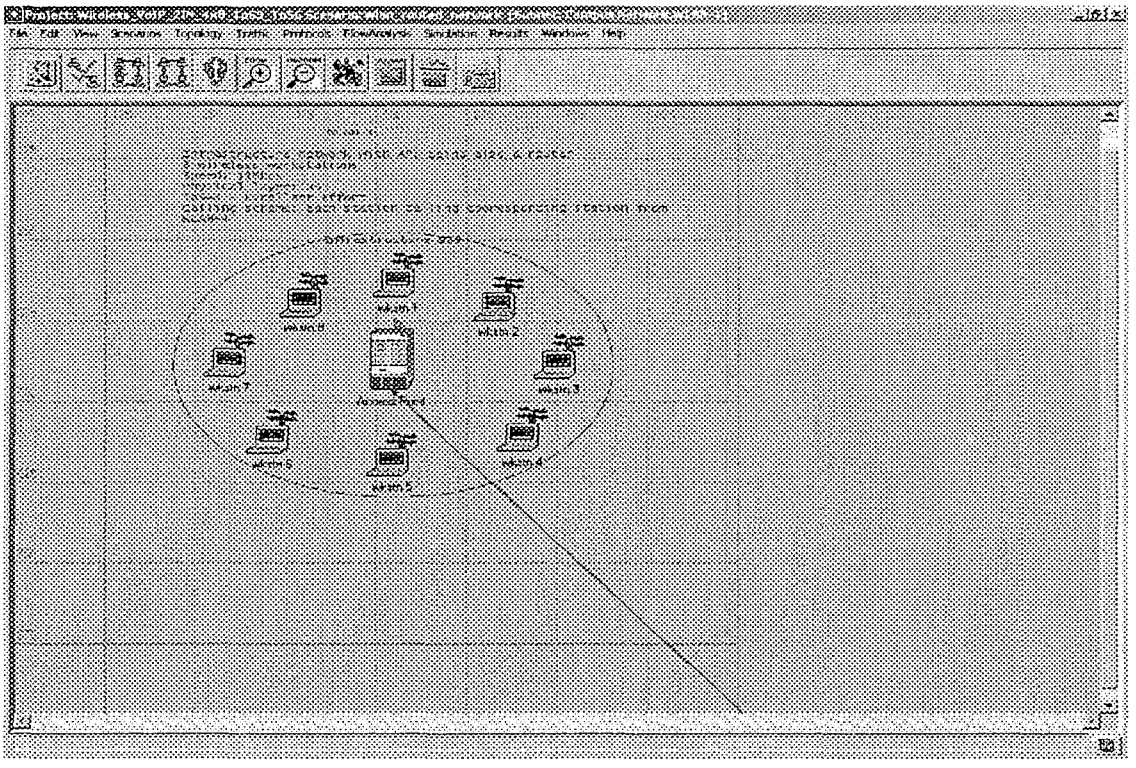
In this subscenario each wireless workstation from WLAN-1 is generating voice calls to the corresponding workstation from WLAN-4 and each wireless workstation from WLAN-2 is generating voice calls to the corresponding workstation from WLAN-3.

- The general architecture for Subscenario 7.2 is:



- The architecture of each of the 4 WLANs is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



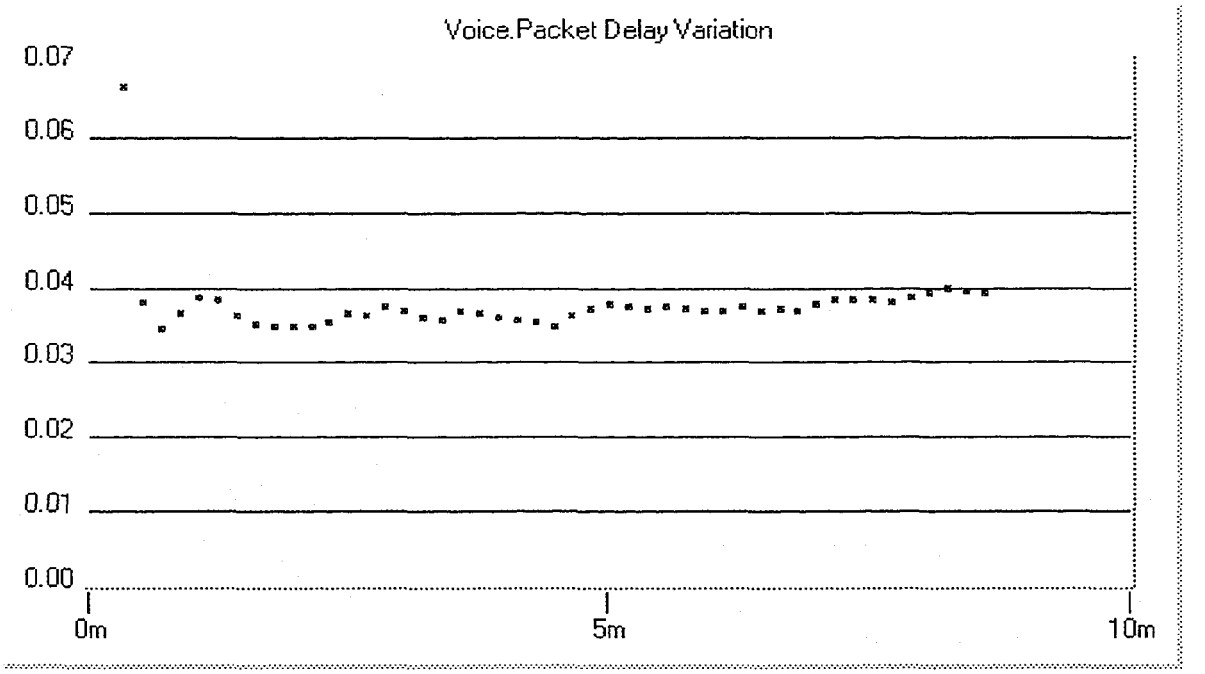
Simulation results are:

- The voice quality parameters are:

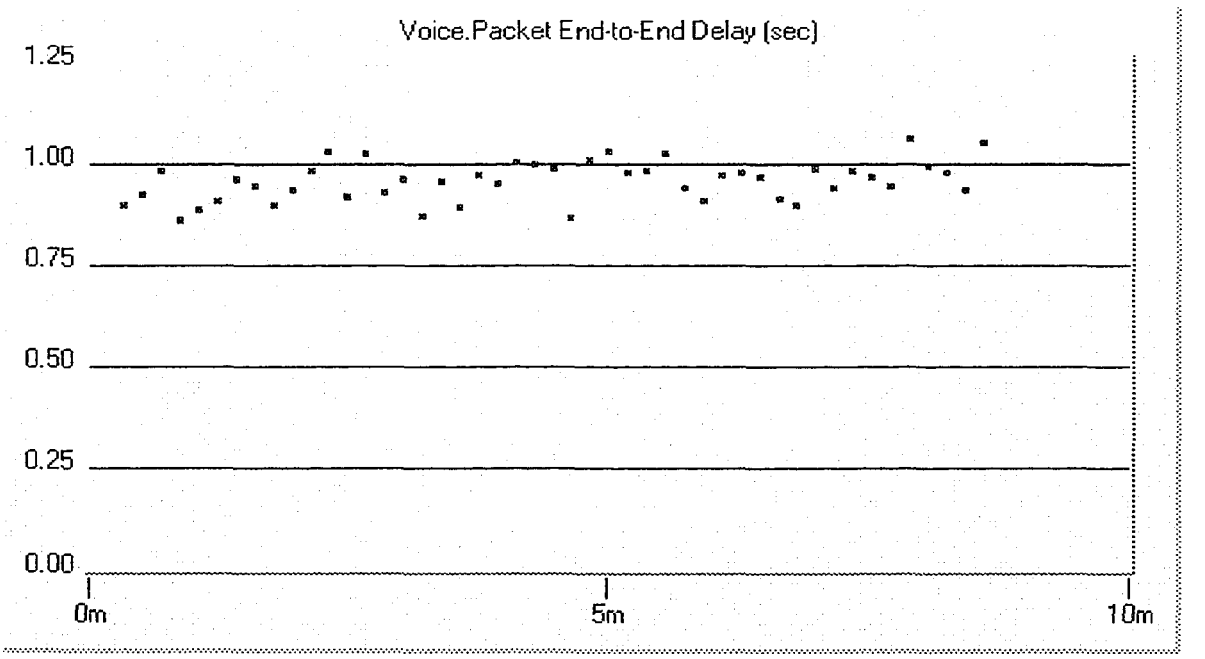
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.0374	0.0666	0.0342
Voice Packet End-to-End Delay (sec)	0.95	1.05	0.85
Voice Traffic Received (bytes/sec)	15,277	17,286	0
Voice Traffic Sent (bytes/sec)	21,547	24,199	0

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Packet delay variation

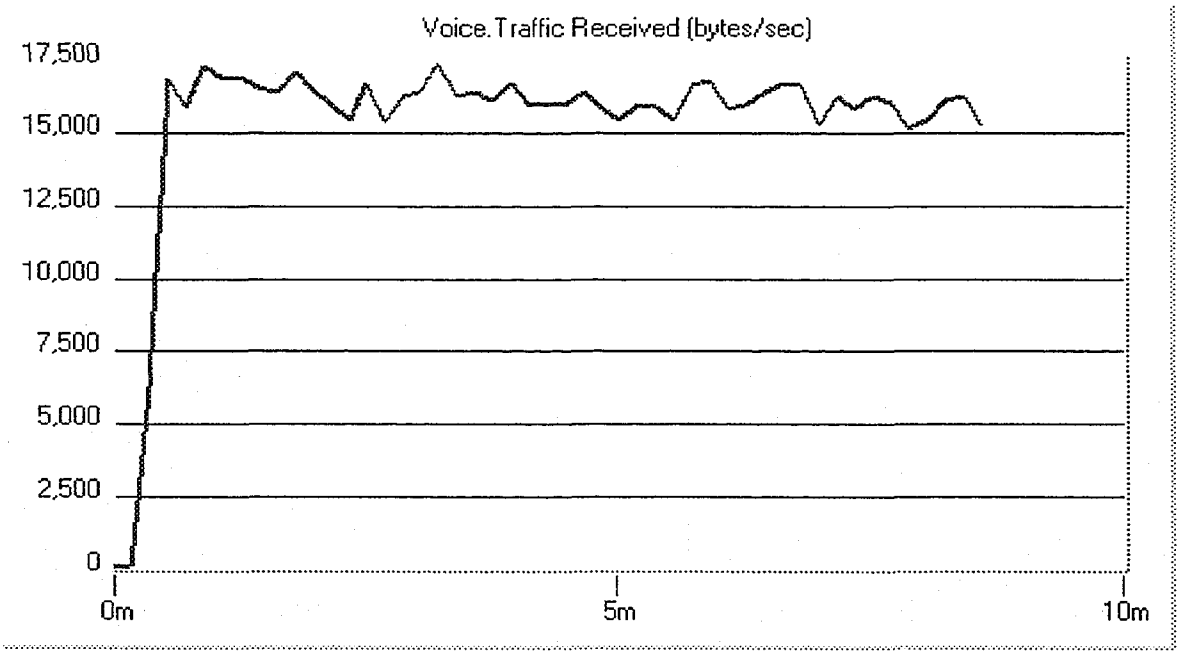


- Graph representation of Voice Packet End-to-End Delay

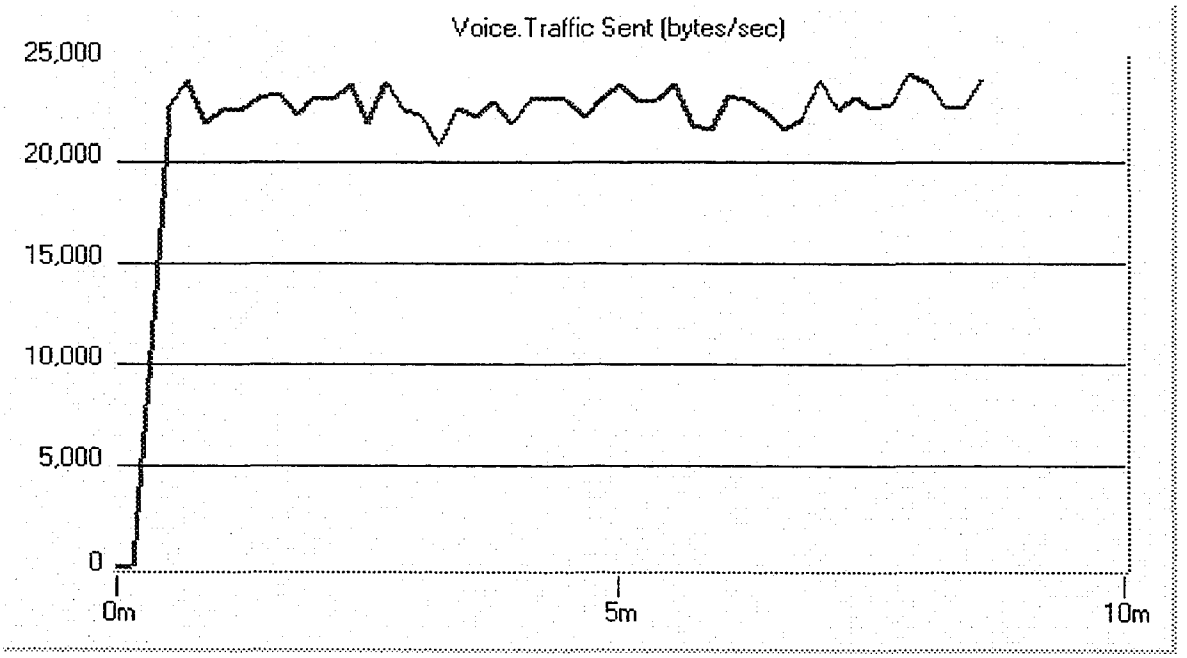


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Traffic Received:



- Graph representation of Voice Traffic Sent:

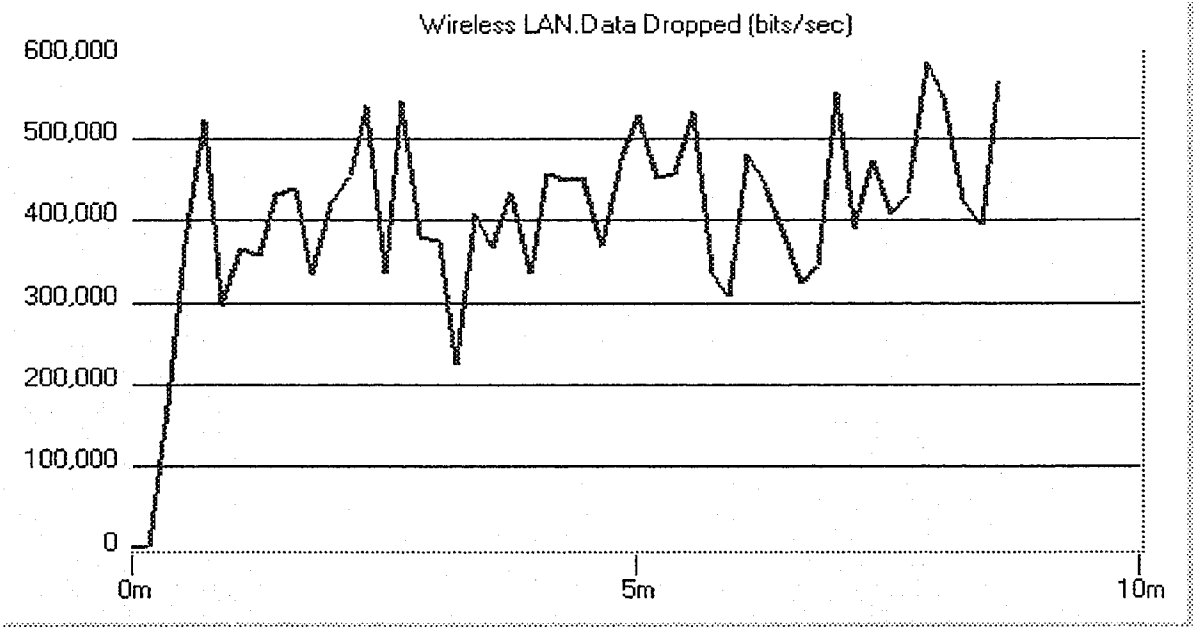


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Wireless Network parameters:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	399,071	587,591	0

- Graphical representation of packet drop rate:



- Comparison table for splitting WLANs:

	2x16 wwkst.	4x8wwkst.
Statistic	Average	Average
Voice Packet Delay Variation	0.00000026	0.0374
Voice Packet End-to-End Delay (sec)	0.000779	0.95
Voice Traffic Received (bytes/sec)	2,601	15,277
Voice Traffic Sent (bytes/sec)	6,605	21,547
Wireless LAN Data Dropped (bits/sec)	13	399,071

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Conclusion for WLAN splitting:

- From the comparison table we can notice that if in case of Subscenario 7.1 the voice quality parameters are still in acceptable range, in case of Subscenario 7.2 they are above the acceptable limits

- Also we notice that in Subscenario 7.2 the number of dropped bits is extremely high. The congestion happens at the routers, which have to route packets between 4 WLANs , instead of only 2.

- So, we can conclude that in case of an architecture with 4 WLANs, interconnected with 2 routers, the 4x8 simultaneous bidirectional calls become are not qualitatively acceptable.

8th Scenario

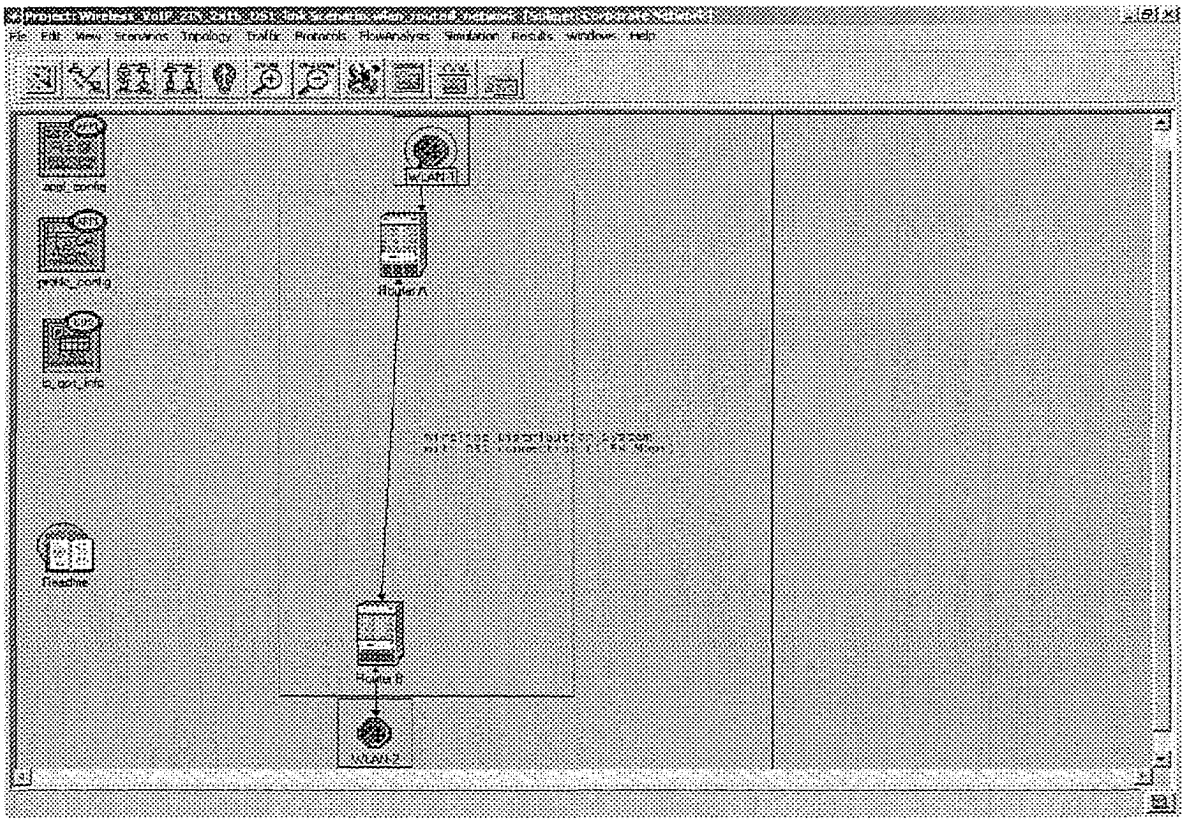
In this scenario we study the impact on the voice quality transmission of a wide area network connection between WLANs.

The network contains two WLANs, locally connected to 1 router each using a 100mbps Ethernet connection, and the two routers are connected with a DS1 line (1.54Mbps).

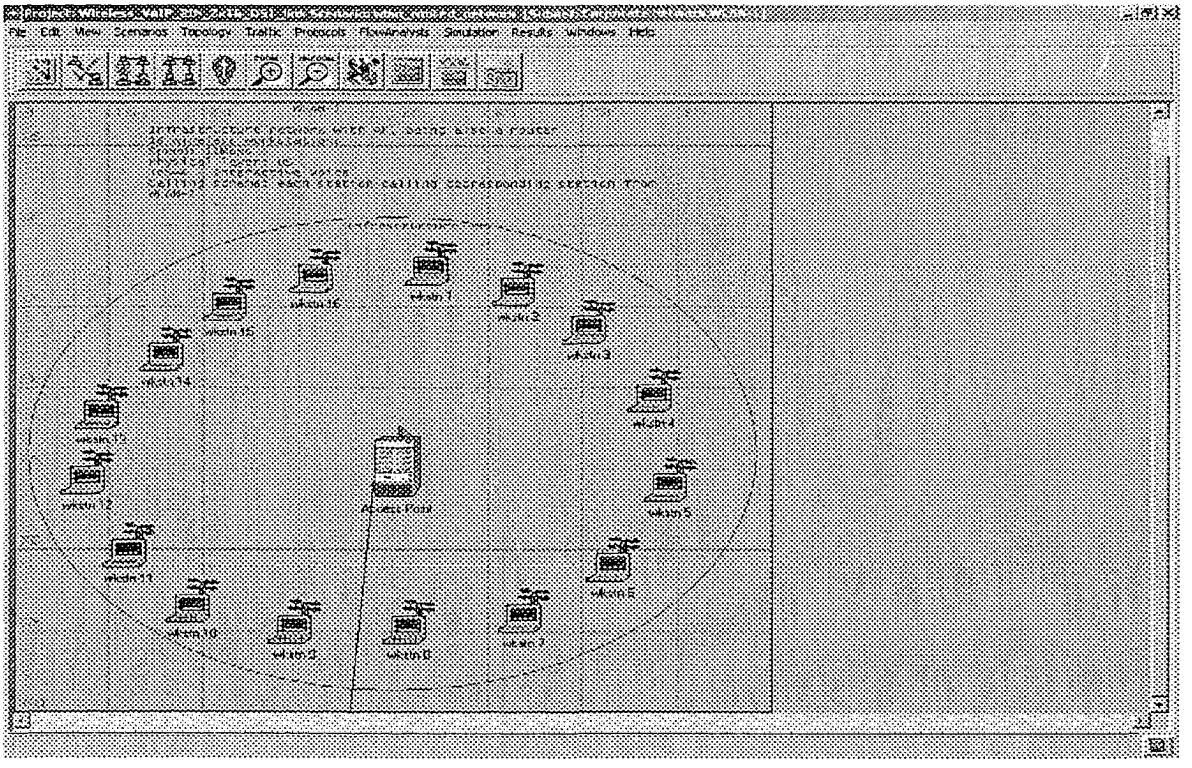
We compare the voice quality parameters from this scenario with the one from Subscenario 6.2 : **WAN connection versus short range connection**.

General Architecture of the network is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- The architecture of each WLAN is:

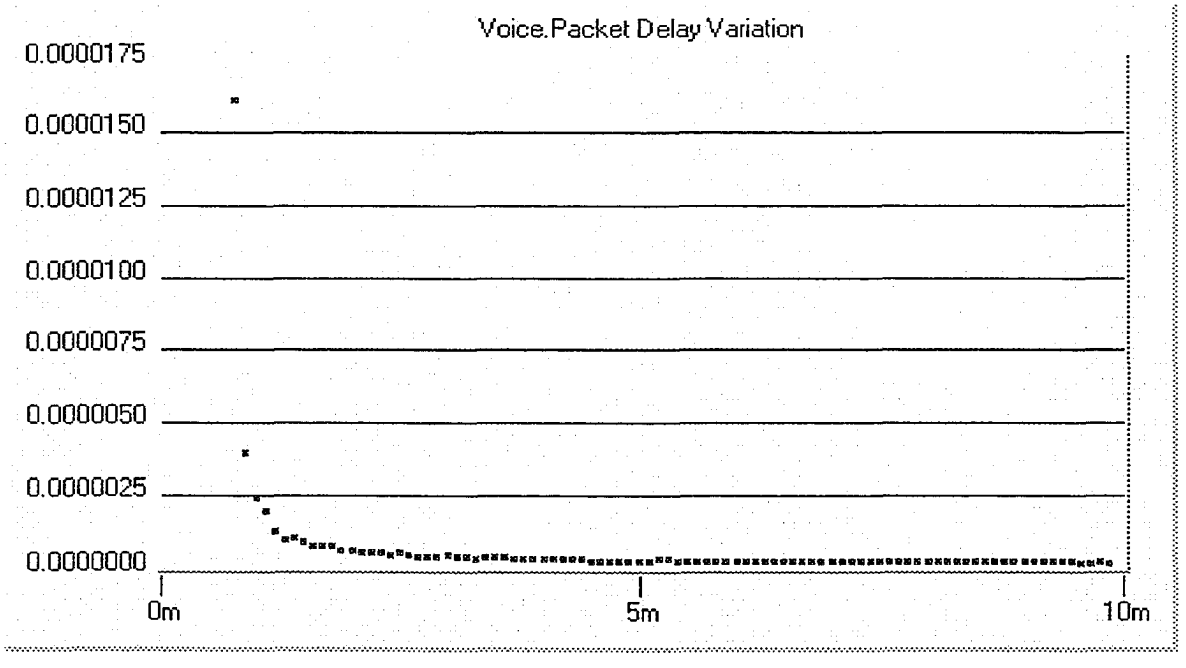


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Voice quality parameters are:

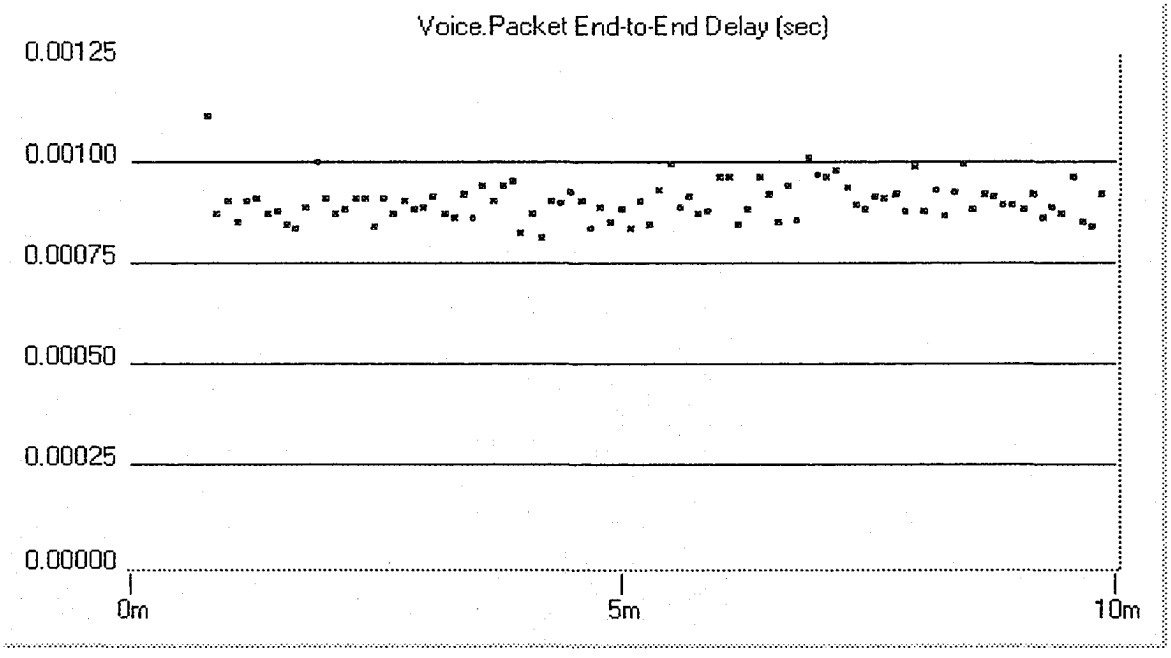
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.0000005	0.0000161	0.0000001
Voice Packet End-to-End Delay (sec)	0.00090	0.00110	0.00081
Voice Traffic Received (bytes/sec)	2,548	3,779	0
Voice Traffic Sent (bytes/sec)	2,549	3,779	0

- Graph representation of Voice Packet Delay Variation:

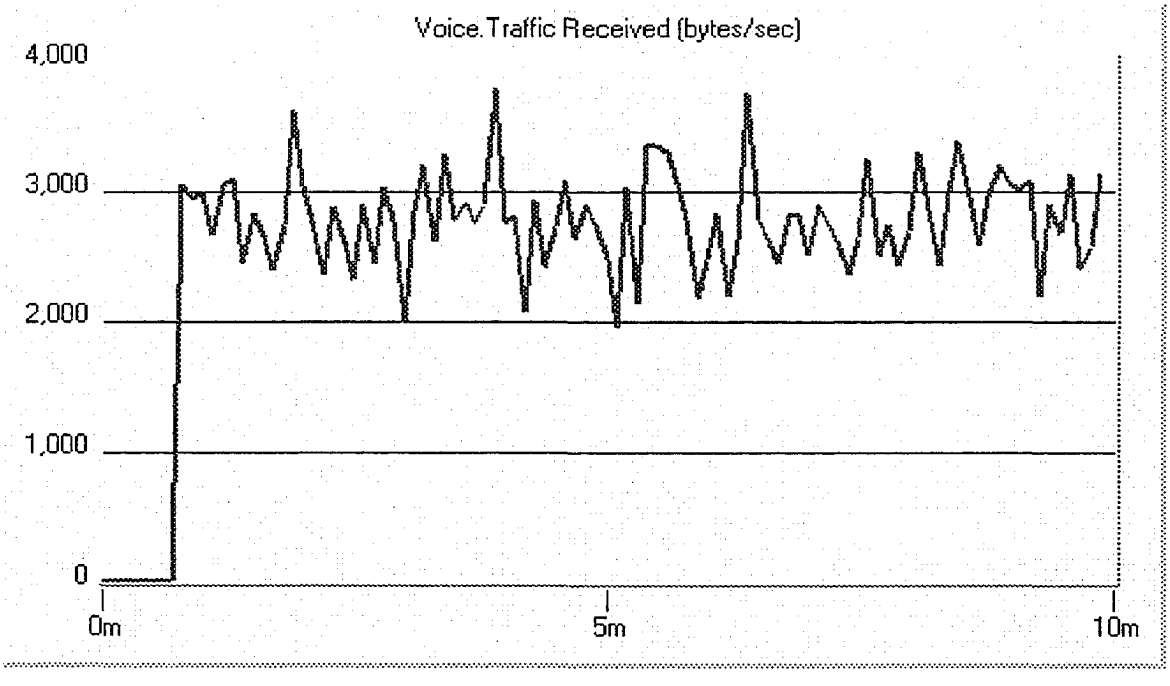


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Packet End-to-End Delay:

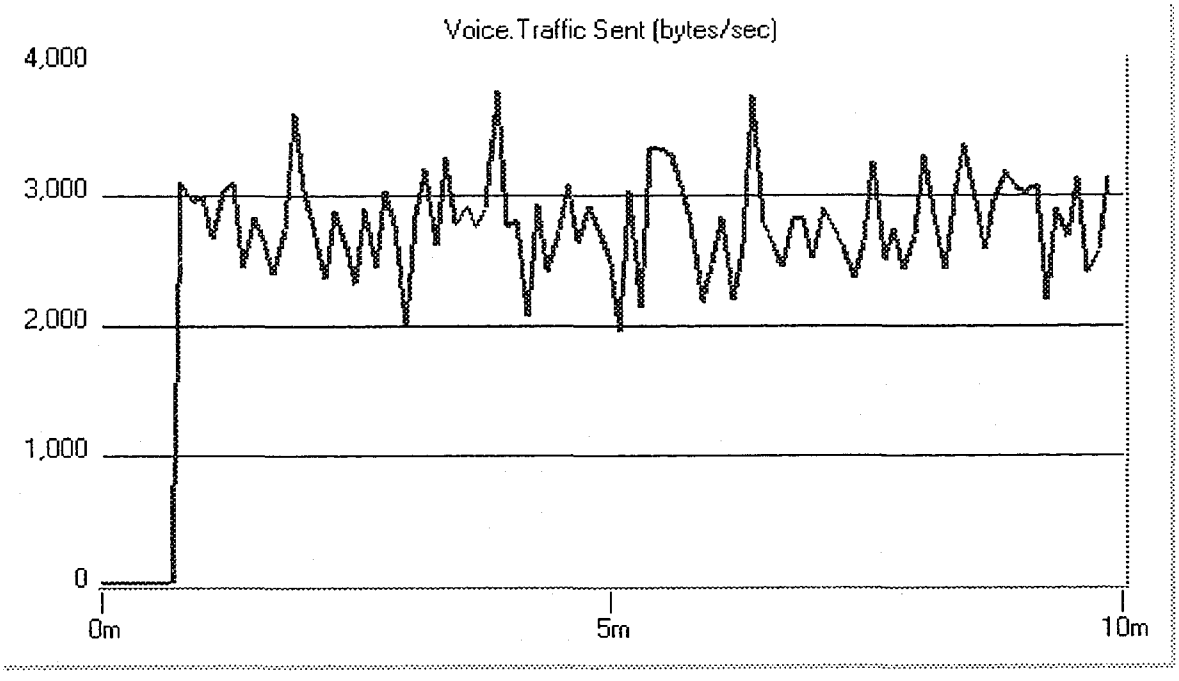


- Graph representation of Voice Traffic Received:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Traffic Sent:



- The relevant parameters of the modeled wireless LAN are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	13	1,344	0
Wireless LAN Delay (sec)	0.000116	0.000230	0.000074
Wireless LAN Load (bits/sec)	326,310	483,755	0
Wireless LAN Media Access Delay (sec)	0.000067	0.000184	0.000000
Wireless LAN Throughput (bits/sec)	326,526	483,755	0

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Comparison table short range vs. long range interconnection:

	802.3 at 100Mbps	DS1 at 1.54 Mbps
Statistic	Average	Average
Voice Packet Delay Variation	0.00000026	0.00000050
Voice Packet End-to-End Delay (sec)	0.000779	0.000900
Voice Traffic Received (bytes/sec)	2,601	2,548
Voice Traffic Sent (bytes/sec)	2,602	2,549

- Conclusion for short distance vs. long distance interconnection:

In the case of 16 of simultaneous calls, comparing a wireline interconnection of 802.3 at 100Mbps, with one of bidirectional DS1 at 1.54 Mbps, the WAN type connection has slightly higher values for both delay and delay variation. Always, the difference is not significant to be able to draw a conclusion.

Note:

1. Logically, the expected result would be to find a drop in the voice performance in case of WAN connection, which did not happen in the case of 16 simultaneous calls.

2. Due to computer power limitation, we were not able to study the same scenarios with higher number of simultaneous calls.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

9th Scenario :

In this scenario we study the impact on the voice quality transmission when WLANs are interconnected through an Internet cloud. An Internet cloud simulates worldwide type connection. The only significant parameter of the Internet cloud is the packet delay and it can be expressed as a constant value or as a probabilistic value. In our simulation we used it as a uniform random variable with a variation in the range of 0.02-0.07 seconds (20 to 70 miliseconds).

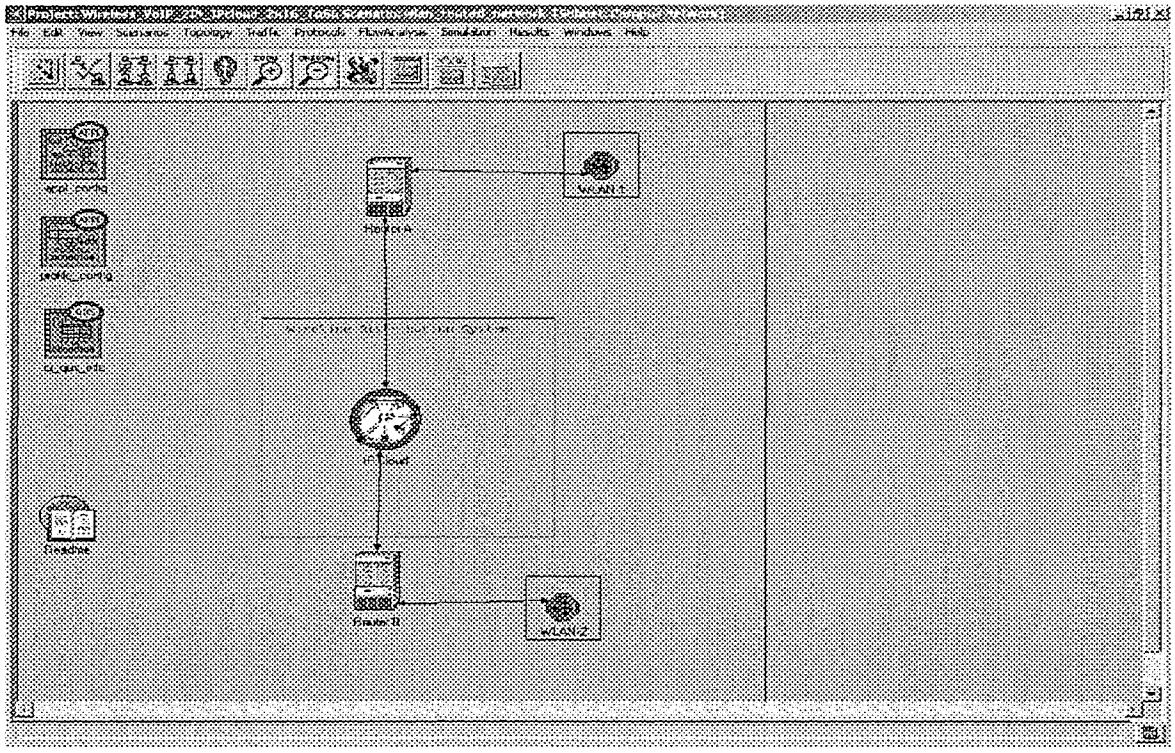
We study the impact on the quality of voice transmission by interconnecting 2 WLANs of 16 workstations each, through an Internet cloud and by interconnecting 4 WLANs of 6 workstations each, through an Internet cloud.

Subscenario 9.1:

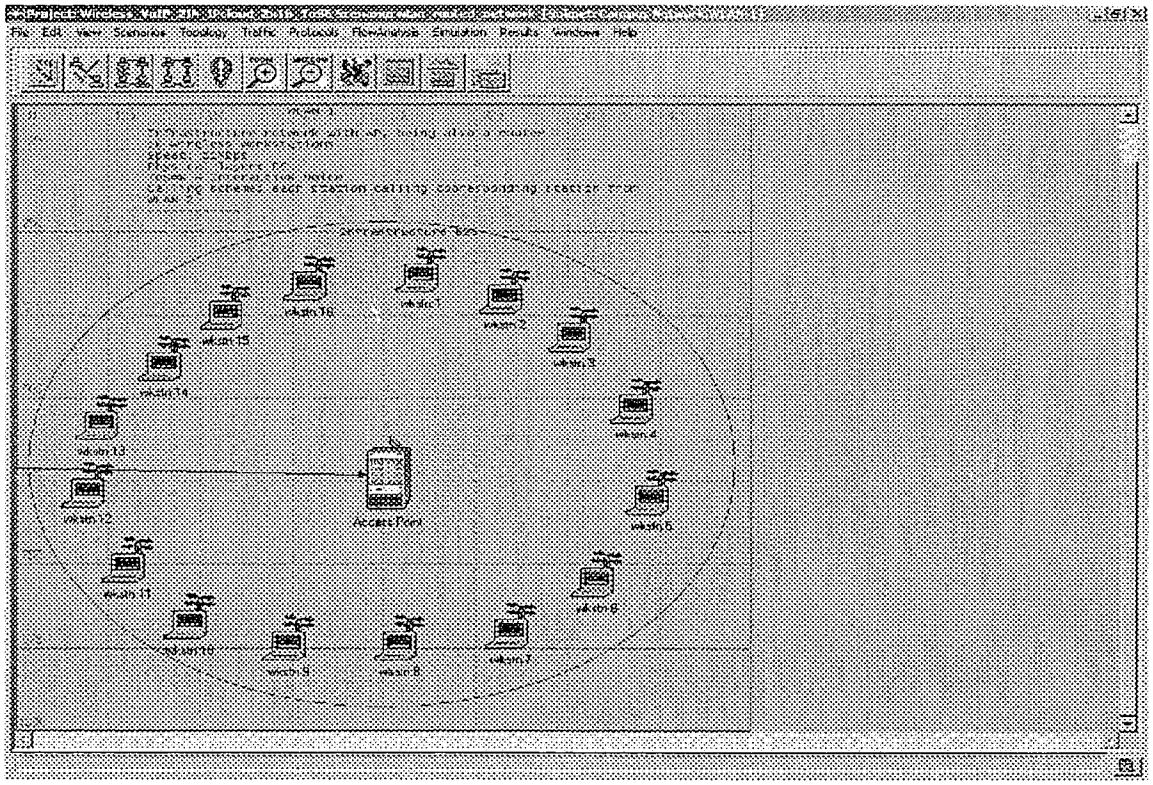
In this sub-scenario we consider 2 WLANs with 16 workstations each, interconnected through 2 routers and an Internet cloud.

- The General Architecture of the network is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- Architecture of each WLAN:

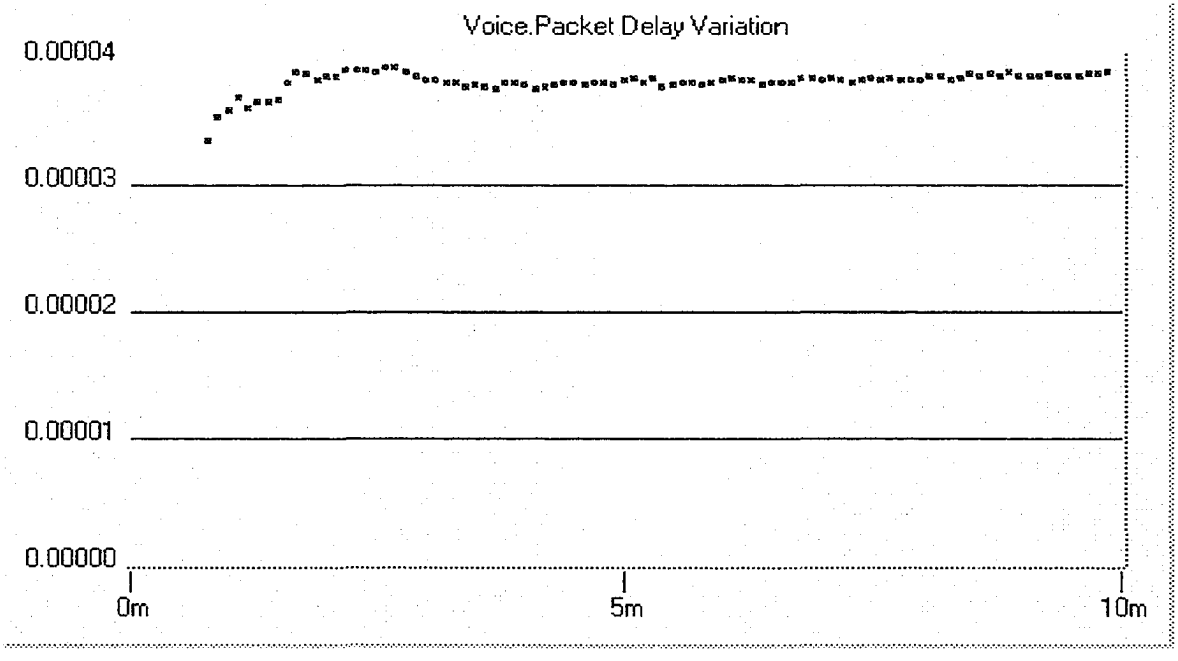


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Voice quality parameters are:

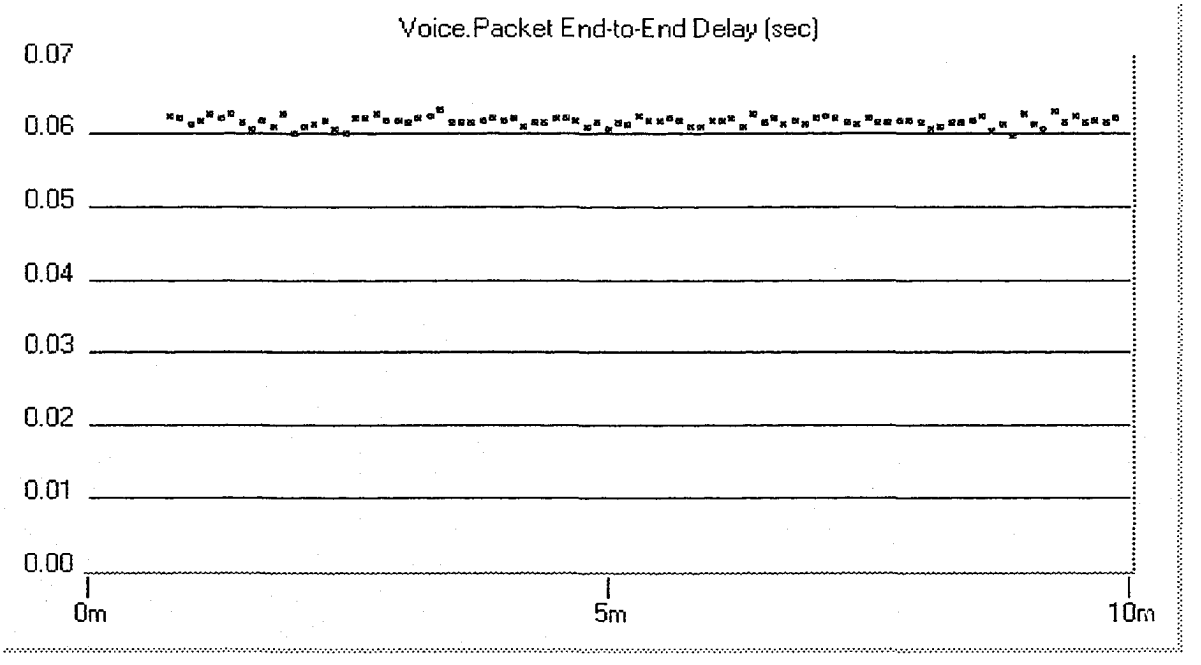
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.0000379	0.0000391	0.0000333
Voice Packet End-to-End Delay (sec)	0.0614	0.0629	0.0595
Voice Traffic Received (bytes/sec)	2,607	3,965	0
Voice Traffic Received (packets/sec)	652	991	0
Voice Traffic Sent (bytes/sec)	2,608	3,921	0
Voice Traffic Sent (packets/sec)	652	980	0

- Graph representation of Voice Packet delay variation:

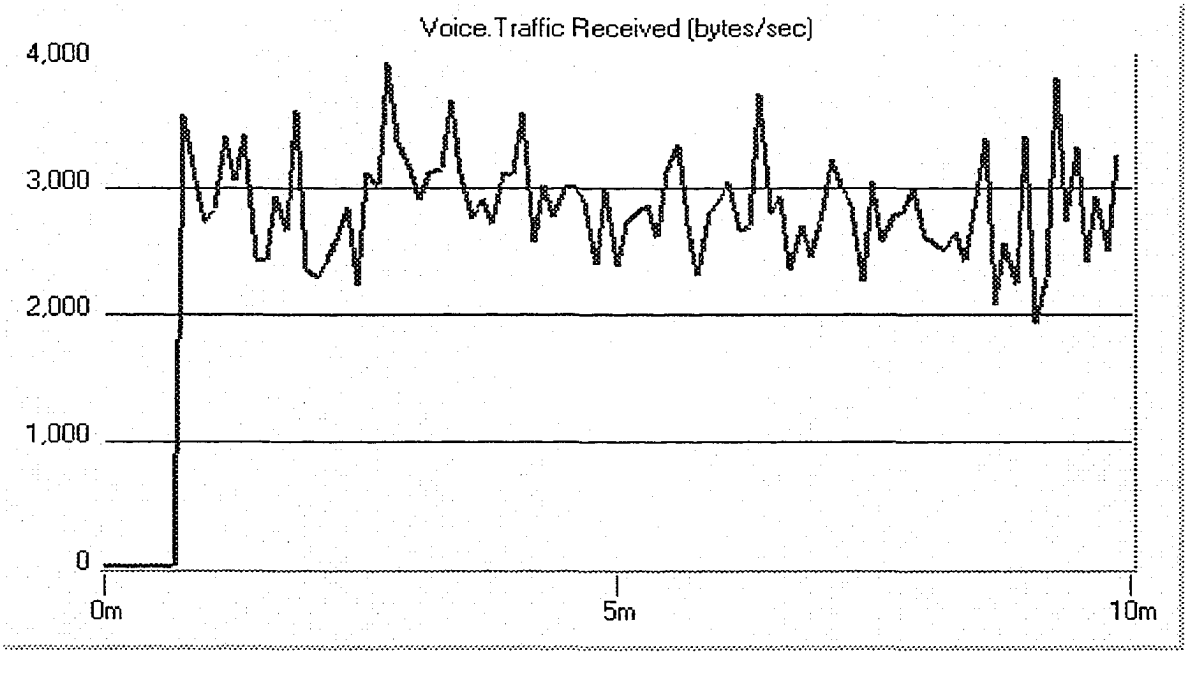


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Packet End-to-End Delay:

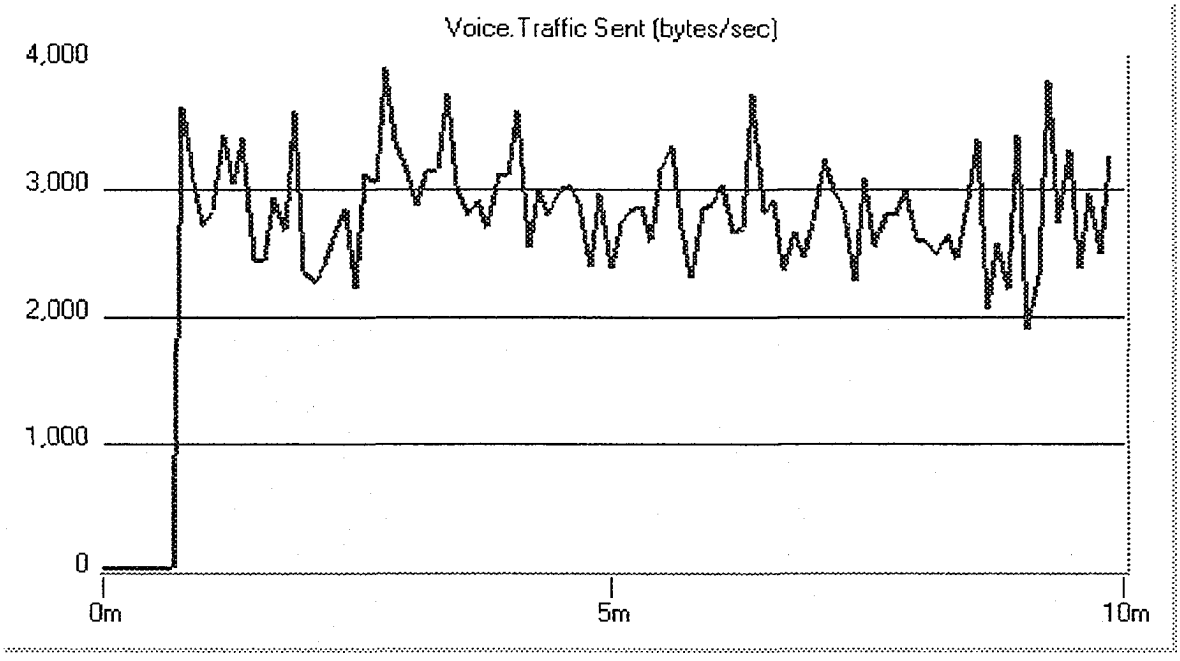


- Graph representation of Voice Traffic Received:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Traffic Sent:



- WLAN traffic parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	13	1,344	0
Wireless LAN Delay (sec)	0.000684	0.000937	0.000088
Wireless LAN Load (bits/sec)	333,941	504,661	0
Wireless LAN Media Access Delay (sec)	0.000631	0.000881	0.000000
Wireless LAN Throughput (bits/sec)	334,200	504,704	0

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Conclusion for Subscenario 9.1:

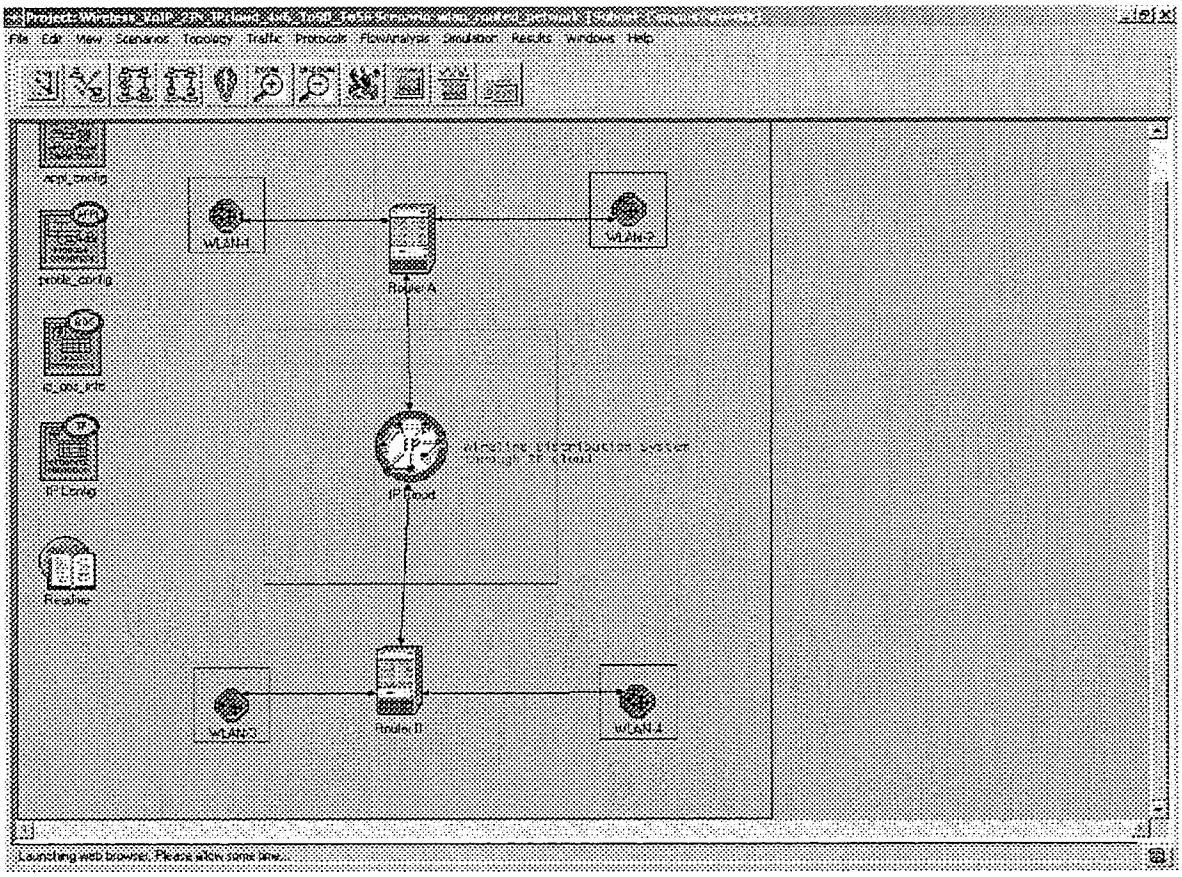
In this simulated architecture we can notice that the average voice packet delay is around 61ms and the delay variation is under 1ms; the average drop rate of 13 bits/sec is also very low. Therefore we can conclude that the modeled and simulated architecture can support quality voice transmission.

Subscenario 9.2:

In this sub-scenario we consider 4 WLANs with 6 workstations each, interconnected through 2 routers and an Internet cloud.

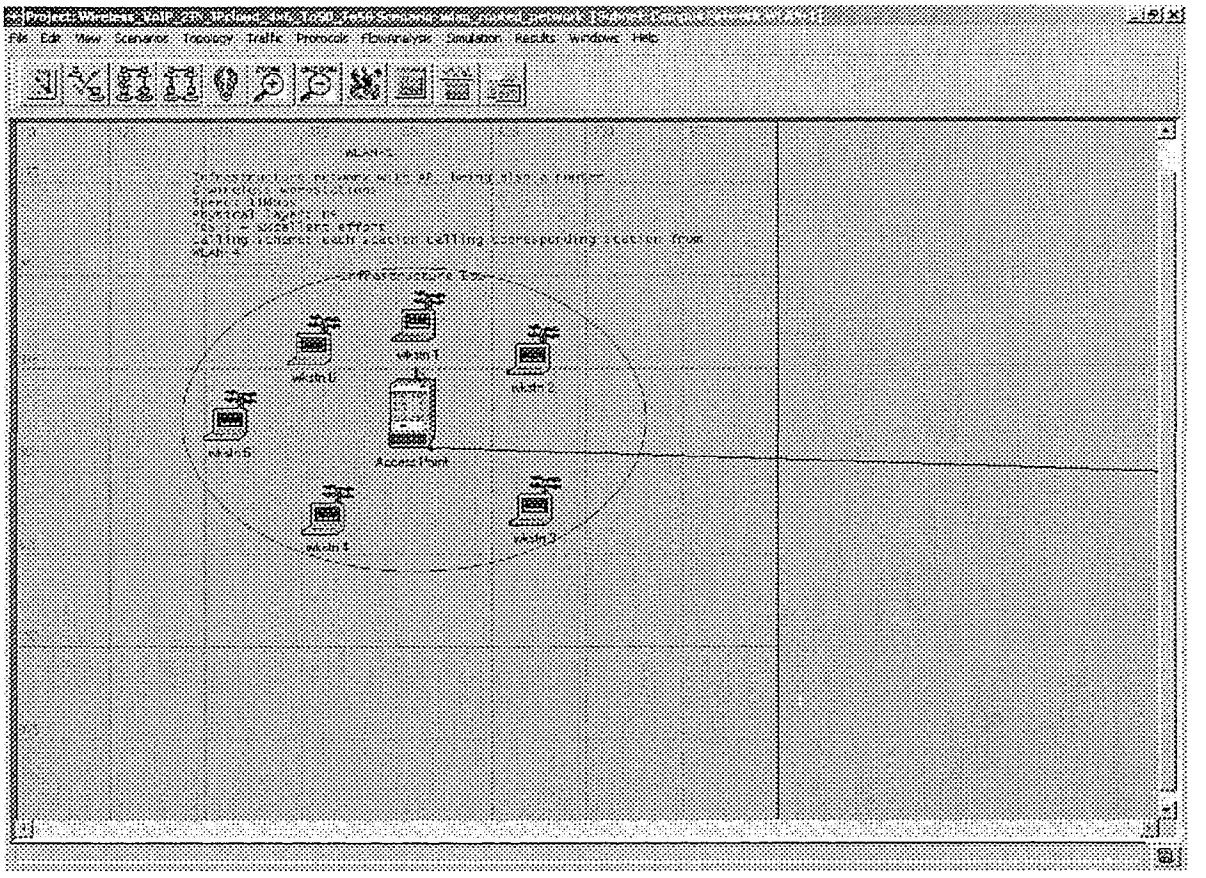
- The general architecture of the network is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- WLAN architecture:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

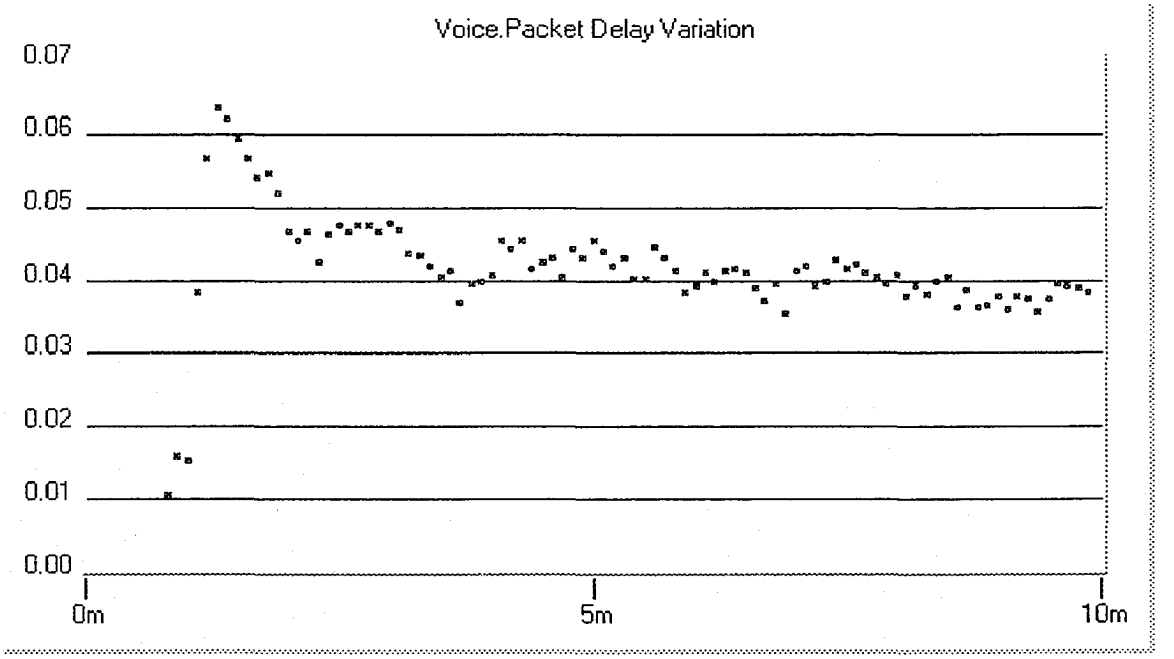


- Voice quality parameters are:

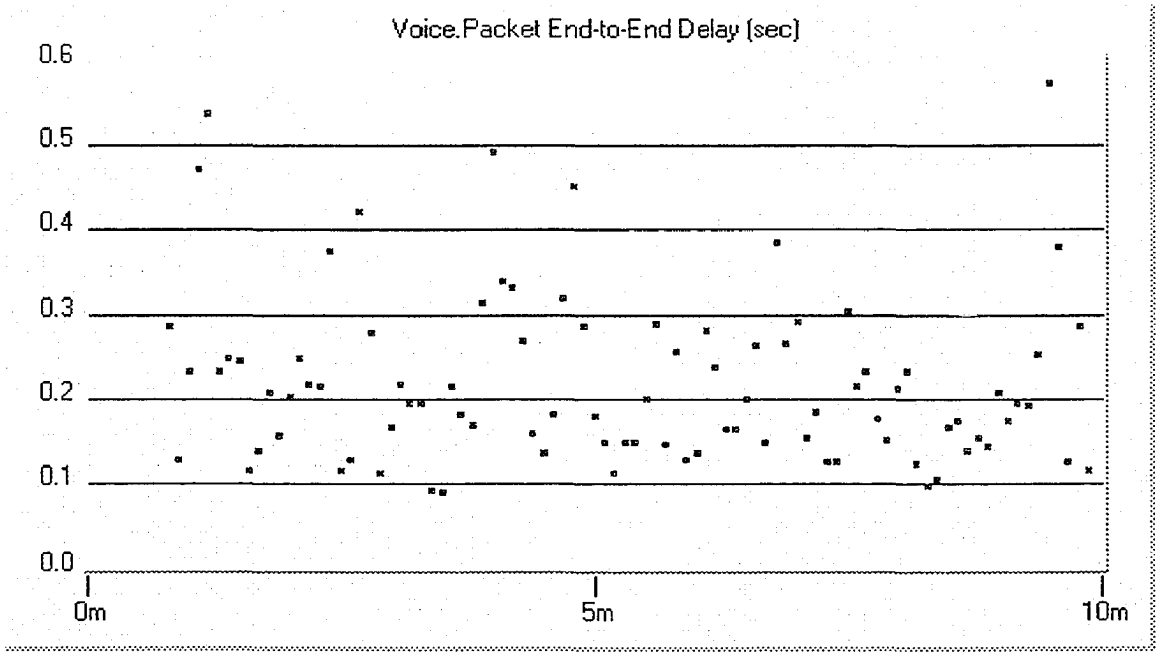
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.0416	0.0633	0.0101
Voice Packet End-to-End Delay (sec)	0.216	0.570	0.088
Voice Traffic Received (bytes/sec)	15,086	17,801	0
Voice Traffic Received (packets/sec)	3,772	4,450	0
Voice Traffic Sent (bytes/sec)	15,158	18,789	0
Voice Traffic Sent (packets/sec)	3,789	4,697	0

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Packet delay variation:

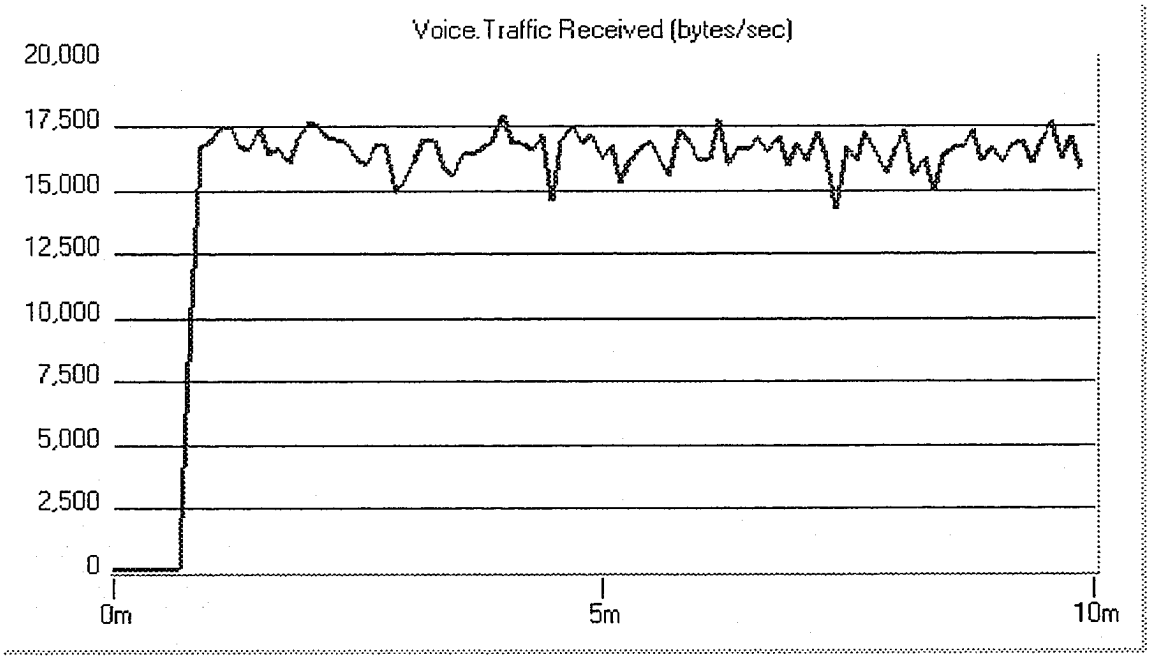


- Graph representation of Voice Packet End-to-End Delay:

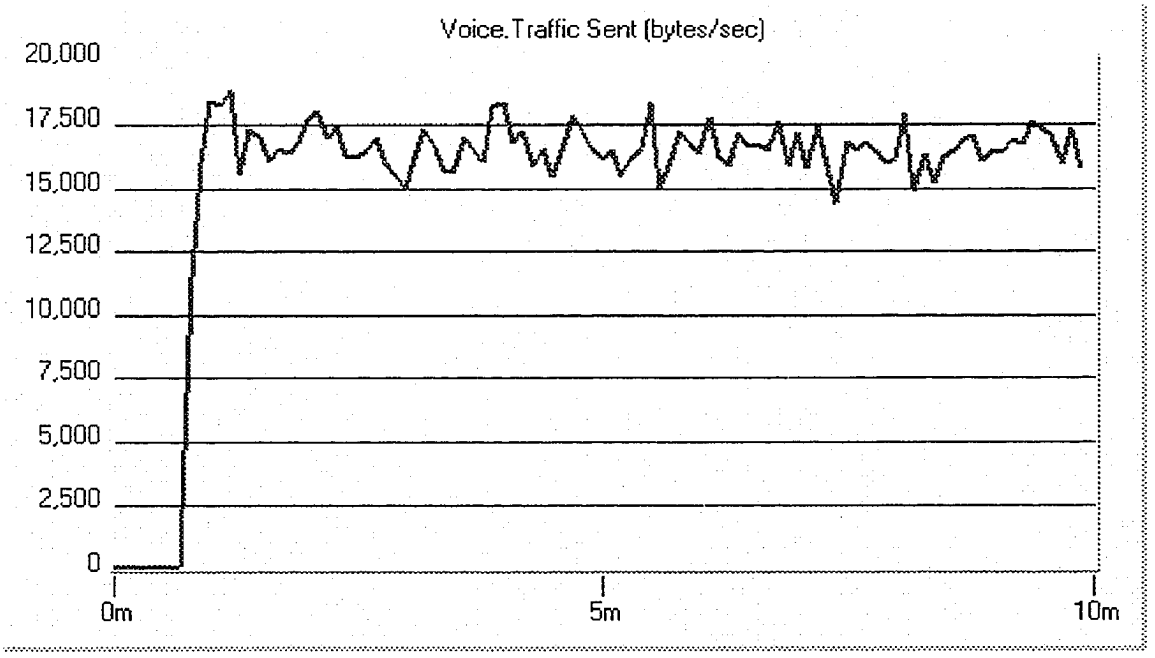


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Traffic Received:



- Graph representation of Voice Traffic Sent:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- WLAN traffic parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	4,111	96,256	0
Wireless LAN Delay (sec)	0.072	0.248	0.000
Wireless LAN Load (bits/sec)	1,935,766	2,331,520	0
Wireless LAN Media Access Delay (sec)	0.071	0.248	0.000
Wireless LAN Throughput (bits/sec)	1,935,814	2,319,659	0

- Conclusion for Subscenario 9.2:

From these results we can conclude, that the simulated architecture drops an unacceptably large amount of packets per second and the voice quality parameters are close to the highest acceptable limit. Due to the large drops rate, this architecture is not suitable for quality voice transmission.

10th Scenario

In this scenario we study the efficiency of the implementation of a Quality of Service in voice traffic routing, by using the "Type of Service" field from the IP protocol, in internetworked WLANs.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Sub-scenario 10.1

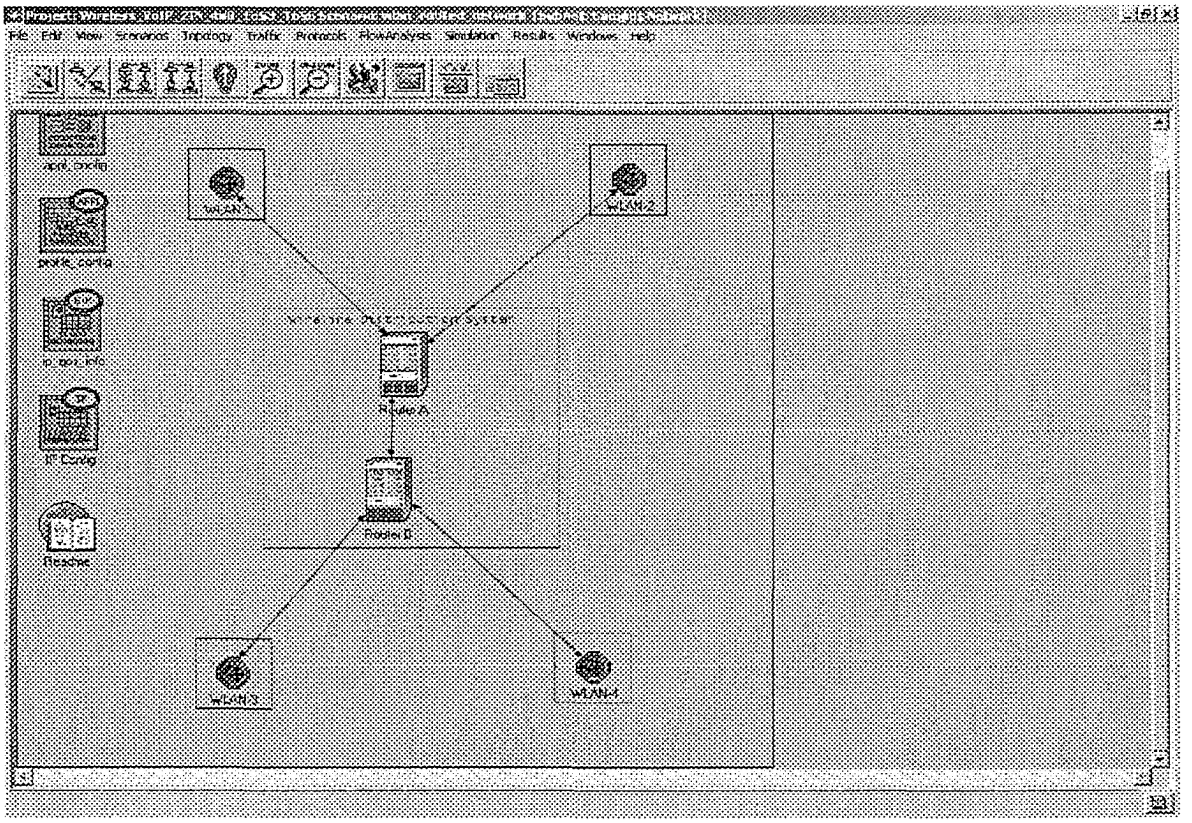
In this sub-scenario we consider 4 WLANs with 8 workstations each, interconnected through 2 routers using a 100Mbps Ethernet connection and the routers are connected using 100Mbps Ethernet connection.

Wireless workstations from WLAN-1 are in bi-directional voice communication with corresponding wireless workstation from WLAN-4 and Wireless workstations from WLAN-2 are in bi-directional voice communication with corresponding wireless workstation from WLAN-3.

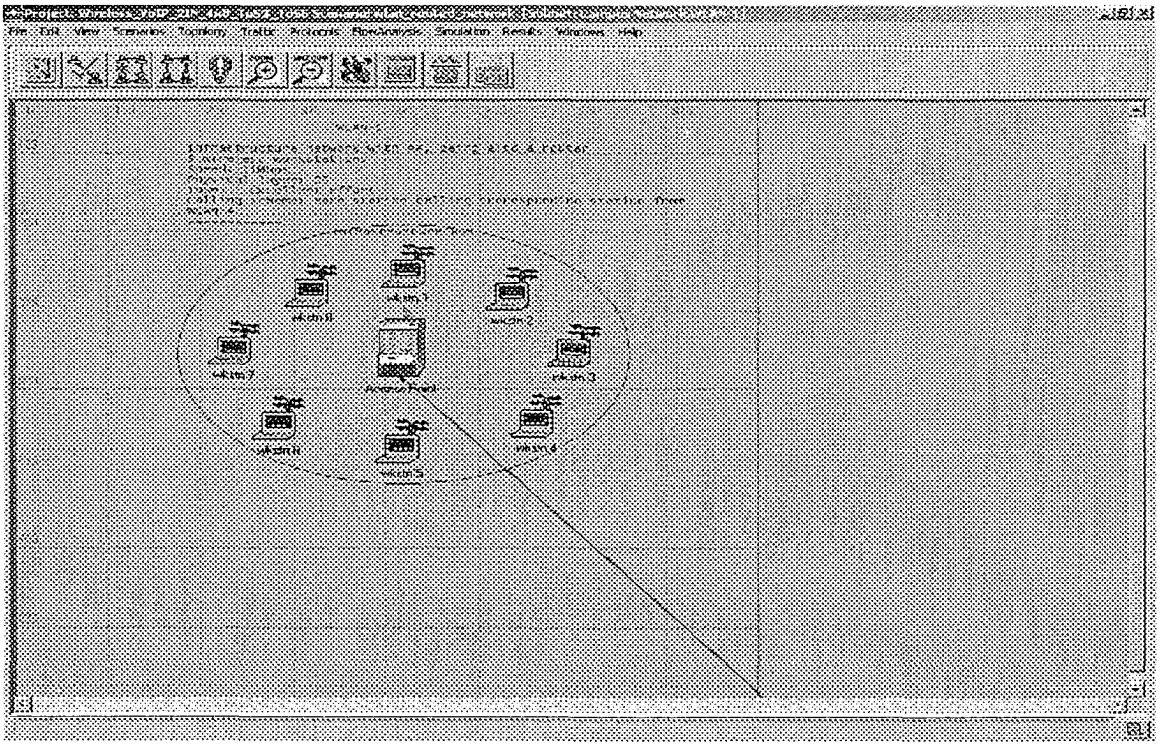
Workstations from WLAN-1 and WLAN-4 are all set with the lowest priority ToS=0 (best effort) and workstations from WLAN-2 and WLAN-3 are all set with the higher priority ToS=6 (interactive voice)

- The general architecture of the network:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



- WLAN Architecture:

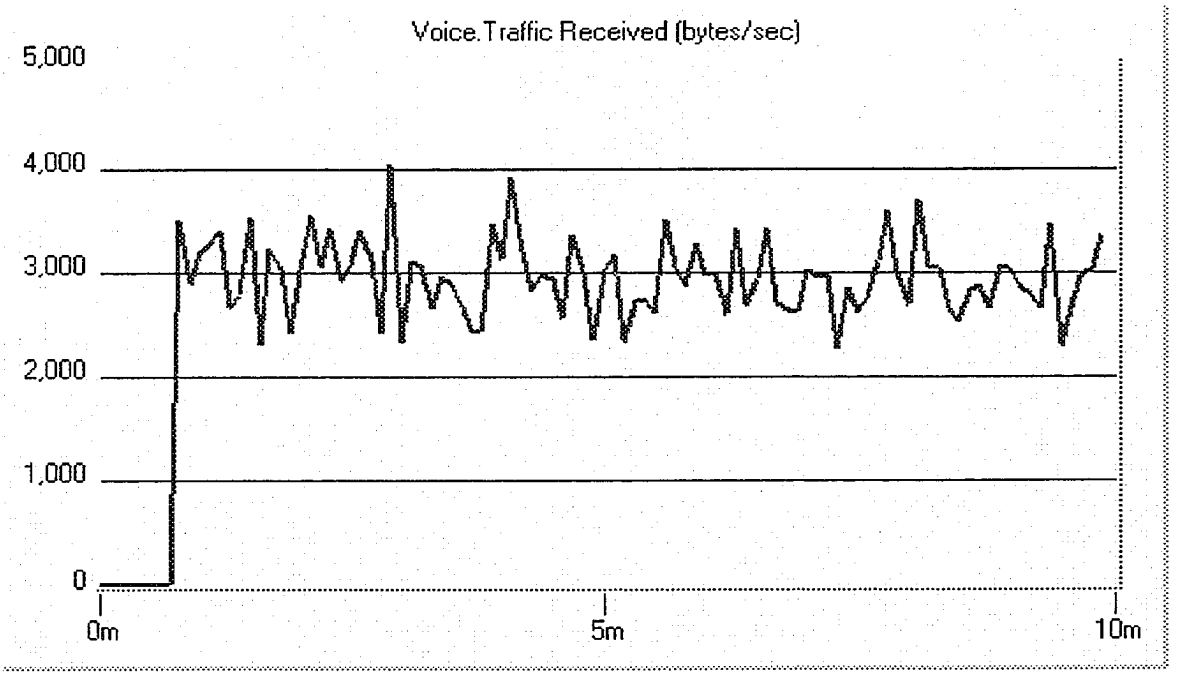


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Voice quality parameters are:

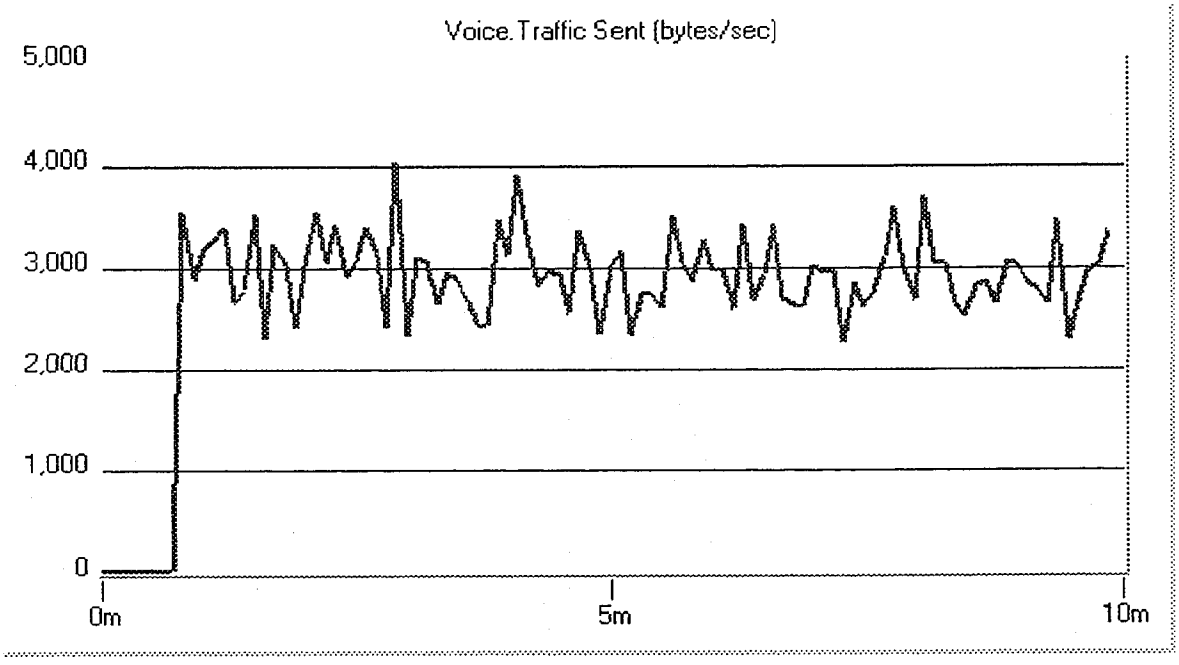
Statistic	Average	Maximum	Minimum
Voice Packet Delay Variation	0.0407	0.0585	0.0352
Voice Packet End-to-End Delay (sec)	0.91	1.02	0.74
Voice Traffic Received (bytes/sec)	12,507	17,870	0
Voice Traffic Received (packets/sec)	3,127	4,467	0
Voice Traffic Sent (bytes/sec)	16,767	23,815	0
Voice Traffic Sent (packets/sec)	4,191	5,954	0

- Graph representation of Voice Packet delay variation:



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Graph representation of Voice Packet Delay:



Note: In order to set up the routing protocol tables and initialize protocol, traffic is delayed for 20 seconds.

- WLAN traffic parameters are:

Statistic	Average	Maximum	Minimum
Wireless LAN Data Dropped (bits/sec)	13	1,344	0
Wireless LAN Delay (sec)	0.000122	0.000246	0.000069
Wireless LAN Load (bits/sec)	345,697	512,811	0
Wireless LAN Media Access Delay (sec)	0.000073	0.000204	0.000000
Wireless LAN Throughput (bits/sec)	345,961	512,853	0

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In this scenario, in addition, we collected the voice quality parameters per workstation, in order to check if wireless workstations from WLANs 2 and 3 outperform those from WLANs 1 and 4.

- Voice quality parameters per workstation:

  - Voice Application Packet Delay Variation

Sort By Node	Sorted By Average	Sort By Peak
WLAN-1.wkstn 8	0.0485	0.0738
WLAN-4.wkstn 3	0.0475	0.0815
WLAN-4.wkstn 1	0.0471	0.0877
WLAN-1.wkstn 2	0.0462	0.0837
WLAN-1.wkstn 5	0.0459	0.0667

  - Voice Application Packet End-to-End Delay (sec)

Sort By Node	Sorted By Average	Sort By Peak
WLAN-1.wkstn 7	0.962	1.12
WLAN-4.wkstn 4	0.961	1.14
WLAN-1.wkstn 8	0.958	1.18
WLAN-1.wkstn 2	0.956	1.15
WLAN-4.wkstn 3	0.955	1.14

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Voice Application Traffic Received (bytes/sec)

Sort By Node	Sorted By Average	Sort By Peak
WLAN-2.wkstn 5	448	935
WLAN-3.wkstn 5	447	799
WLAN-2.wkstn 4	443	878
WLAN-3.wkstn 4	434	836
WLAN-2.wkstn 8	431	803

- Voice Application Traffic Sent (bytes/sec)

Sort By Node	Sorted By Average	Sort By Peak
WLAN-2.wkstn 5	572	1,065
WLAN-3.wkstn 5	570	1,025
WLAN-1.wkstn 2	568	977
WLAN-4.wkstn 5	566	975
WLAN-3.wkstn 4	563	974

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Conclusion for QoS implementation:

We notice that in both delay and delay variation, per workstation, the top workstations belong to WLAN-1 and WLAN-4, were the Quality of Service parameter is ToS=0 (best effort).

The voice traffic is symmetrical and comparable in value for both flows of voice traffic. This proves that at the router interface, the packets coming from WLAN-2 and WLAN-3, with type of service 6 (interactive voice), are given higher priority over packets coming from WLAN1 and WLAN-4, with type of service 0 (best effort).

So, we can conclude that "Type of Service" field from the IP protocol, can be used to prioritize voice call, in case of internetworked WLANs, when packet routing is necessary.

Bibliography

13.1. Leung, Kin K., B. McNair, L. Cimini, J.H. Winters. "Outdoor IEEE802.11 Cellular Networks: MAC Protocol Design and Performance". IEEE Communications Magazine. pg. 595-5999. Feb. 2002.

13.2. Lucas, Matthew T., B.J. Dempsey, D.E. Wrege, A.C. Weaver. "An Efficient background Traffic Model for Wide-Area Network Simulation". Proceedings of Sixth International Conference on Computer Communications and Networks. Sept. 1997.

13.3. Motorola. Forrester, Nigel. "Measuring Performance and Quality of packet Voice Systems", Whitepaper. Jan. 2003.

13.4. Nacht Telecommunications Inc. "VoIP Industrial tutorial". April 11, 2001.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

13.5. Ramakrishna, Sudhir, J. M. Holtzman. "A Scheme for Throughput Maximization in a Dual-Class CDMA System". IEEE Journal on Selected Areas in Communications, Vol. 16, No. 6, August 1998.

13.6. Sanneck, Henning. N.T. Long Le, M. Hardt, W. Mohr. "Selective Packet Prioritization for Wireless Voice over IP".

13.7. Veeraraghavan, N., N. Cocker, T. Moors. "Support of Voce Services in IEEE.802.11". Proceedings IEEE INFOCOM, Anchorage, AK. April 2001.

13.8. Sanneck, H., N. Le, A. Wolisz, G. Carle. "Interflow Loss Recovery and control for VoIP". Proceedings ACM Multimedia, Ottawa, Canada. Sept. 2001.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Conclusions

We made a general presentation of the evolution of wireless communications. We emphasized on the need of convergence of different type of communications: voice, video and data.

Further we studied the performance of the voice transmission over new generations of wireless networks.

With this in mind, we conceived different wireless architectures and studied through modeling and simulation, the main performance parameters of the voice quality and of the network.

Three types of wireless network architectures are considered:

- Ad-hoc wireless LAN,
- Infrastructure (named also backbone) wireless LAN
- Wireless LANs interconnected with wireline routers.

In all scenarios 802.11b WIFI wireless communication protocol is simulated. In case of internetworking, UDP over IP is used as transport layer level protocol.

We generated digitized voice over packet switching network.

The following voice quality parameters are analyzed:

- Voice packet delay variation (jitter)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Voice packet end-to end delay
- Voice traffic sent (bytes/sec)
- Voice traffic received (bytes/sec)
- Packet drop rate (packets/sec)

The performance parameters are analyzed at global network level or if necessary, at individual wireless workstation level.

1. We were able to determine that G.729 codec outperforms G.711 in WLAN voice communication
2. Also we determined that at physical layer level, Direct Sequence outperforms Frequency Hopping for voice communication.
3. We were able to see a small negative impact on the quality of voice transmission in case of a limited mobility in the communication area.
4. We were able to determine the impact of the number of wireless users over the quality of the voice transmission: as the umber of users increases the voice quality deteriorates. Due to limited resources in PC power, we were not be able to determine the number of simultaneous users that make the quality of the voice transmission unacceptable.
5. We studied the impact on the quality of the voice transmission in different types of internetworking: campus wide, corporate wide and worldwide.
6. Finally we studied the implementation of the Quality of service (QoS) in voice transmission over intrenetworked WLANs, by using the field "Type of Service" from the IP protocol; and we found it as a first acceptable step in QoS implementation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The study can be continued at least in the following directions:

1. The impact of other web applications like: e-mail, telnet, web browsing etc. over the quality of voice transmission

2. The implementation of more efficient Quality of Service schemes than is the use of the "Type of Service" filed from the IP protocol.

3. The combination of more than one internetworking architectures over the quality of the voice transmission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix A. Queuing Theory

The theoretical basis of communications networks is the queuing theory.

Queues are essential to network systems. We encounter a large variety of queues in a communication network:

- queues of inquiries waiting to be processed by a network element, queue of data base requests, queues of I/O requests, etc.
- In circuit switching: a call traverses several nodes in the network and each node has a queue for calls requesting a channel.
- In packet switching: routers have dynamic route allocation and routing is dependent on the length of the queue of packets.

The main components of a queue are:

- Arrival process: is the probability density distribution that determines the customer arrivals in the system
- Service process: is the probability density distribution that determines the customer service times in the system
- Number of Servers: available to service the customers.

The process can be described as follows:

1. customers of a certain population enter the queuing system to get a certain service.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Customer has a generic meaning here; for example, in case of a communication network messages or packets are the customers.

2. Upon arrival, a customer is either served, if a server is available, or sits in a waiting room; usually a buffer.

3. After being served, the customer leaves the queuing system.

Kendall introduced a notation in order to symbolize a queuing system; the generic form of the notation is:

A/S/n/C where:

A describes the interarrival time distribution

S the service time distribution

n the number of servers

C the size of the system capacity, which include the servers.

For A and B it became usual to get the values:

M for exponential distribution (M stands for Markov)

D for deterministic distribution (all customers have the same value)

G (or GI) for any type of probability distribution.

Some of the widely known queuing systems are:

M/M/1: Poisson arrivals, exponentially distributed service times, one server,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

infinite buffer

M/M/m: same as previous with m servers

M/M/m/m: Poisson arrivals, exponentially distributed service times, m server, no buffering

M/G/1: Poisson arrivals, identically distributed service times follows a general distribution, one server, infinite buffer

*/D/∞: A constant delay system

In order to be able characterize the above mentioned queuing systems, we introduce first the concept of Markov process.

A.1 Markov Process

A Markov process is defined in the state-time space.

Therefore, one can distinguish continuous state Markov processes and discrete state Markov processes.

Discrete state Markov processes are also named Markov chains.

A Markov chain could be also continuous or discrete in time.

Of a paramount importance for communication networks are the discrete time Markov chains, which could be simply named Markov chains.

The general definition of a Markov process is:

a Markov process is a stochastic process $(X(t), t \in T)$, $X(t) \in E \subset R$, such that

$$P(X(t) \leq x \mid X(t_1) = x_1, \ldots, X(t_n) = x_n) = P(X(t) \leq x \mid X(t_n) = x_n)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

for all $x_1, \ldots x_n$, $x \in E$, and $t_1, \ldots, t_n, t \in T$, with $t_1 < t_2 < \ldots < t_n < t$.

Intuitively, the formula above says that the future of the process depends only on the current state and not on the history of the process. Or, in other words, the entire history of the process is summarized in the current state.

A.2 Birth-Death Process

Next we introduce the concept of birth/death process:

A birth and death process $(X(t), t \geq 0)$, is a continuous-time, discrete-space Markov process, with the following properties:

1. $P(X(t+h) = n+1 \mid X(t) = n) = \lambda_n h + o(h)$ for $n \geq 0$

2. $P(X(t+h) = n-1 \mid X(t) = n) = \mu_n h + o(h)$ for $n \geq 1$

3. $P(X(t+h) = n \mid X(t) = n) = 1 - (\lambda_n + \mu_n)h + o(h)$ for $n \geq 0$

The random variable X(t) can be understood as being the size of the population at a time t.

λ_n represents the birth rate, and

μ_n represents the death rate,

when the size of the population is n.

It can be demonstrated that in a birth-death process, the following balance equations are true:

If we note:

$P_n(t) = P(X(t) = n)$ for $n \in N, t \geq 0$ and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$\text{Lim}_{t->\infty} P_n(t) = p_n$, then we have:

$\lambda_0 p_0 = \mu_1 p_1$

...

$(\lambda_n + \mu_n) p_n = \lambda_{n-1} p_{n-1} + \mu_{n+1} p_{n+1}$, $n=1,2,3, \ldots$

These balance equations express the fact that the probability of a flow entering a state is equal with the probability of the flow exiting that state.

A.3 Little's formula

In a queuing system, performance issues can be addressed from two different point of view: the system point of view and the user point of view.

From system point of view, performance issues are like: the limiting distribution of the queue length, the mean number of customers, etc.

From the user point of view, performance issues are like: response times and waiting times.

One of the most useful result in performance evaluation is Little's formula, presented below.

Let's consider the following notations:

- $\lambda \geq 0$ is the arrival rate of the customers in the queue
- $\underline{N}$ is mean number of customers in the system, and
- $\underline{T}$ is the mean total waiting time of a customer in the system (T includes both the proper waiting time as well as the serving time).

Using these notations, we have:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$\underline{N} = \lambda \underline{T}$

In words, Little's formula states that the average number of customers in a queuing system, which is in steady-state, is linearly proportional with the arrival rate in the system and the average time spent in the system.

Intuitively, we can expect that the system being in steady-state, the arrival rate λ to be equal with the departure rate $\underline{N}/\underline{T}$.

A.4 Queuing systems

Further, we present the main characteristics of some of the well-known queuing systems.

1. The M/M/1 Queue

By definition the M/M/1 queuing process has:

- the interarrival time distribution is a Markov chain (Poisson arrivals)
- the service time distribution is also a Markov chain (Exponential service time)
- the number of servers is 1
- the buffer length is infinite.

In this queuing system the customers arrive according to a Poisson process, with rate λ. The time it takes to serve every customer is an exponential random variable with parameter μ., named also exponential service times.

The service times are independent one from the other and independent of the interarrival times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

If a customer enters an empty queue, she/he is served at once; if the system is not empty then will join the queue.

Let X(t) be the number of customers in the system at time t.

It is provable that the process (X(t), $t \geq 0$) is a birth and death process with birth rate

$\lambda_i = \lambda$, for all $i \geq 0$, and with death rate

$\mu_i = \mu$ for all $i \geq 1$.

Let $\pi(i)$, $i \geq 0$, be the density function of the number of customers in the system in steady-state.

The balance equations for this birth and death process are:

$\lambda \pi(0) = \mu\pi(1)$

$(\lambda + \mu) \pi(i) = \lambda\pi(i - 1) + \mu\pi(i + 1)$ for any $i \geq 1$.

The notation:

$\rho = \lambda / \mu$

is named the traffic intensity of the system, because it expresses the mean quantity of work per unit of time, brought into the system.

One can prove that, if $\rho < 1$, then

$\pi(i) = (1 - \rho) \rho^i$ for any $i \geq 0$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This equation expresses the fact that, in case of $\rho < 1$, the system is stable, if the work that is brought to the system per unit of time is strictly smaller than the processing rate, which in our case is 1, as we have only one server.

In this case, the mean number of customers E[X] is:

$$E[X] = \frac{\rho}{1-\rho}$$

We can notice two things:

$\lim_{\rho \to \infty} E[X] = \infty$, which practically says that, if the system is not stable, then the queue will be infinitely big.

$\pi(0) = 1 - \rho$, which means that, if the system is stable, then the queue will empty infinitely many times.

The probability for the queue to exceed k customers, in steady-state, is:

$P(X \geq K) = \rho^k$

The throughput T, of an M/M/1 queue in equilibrium, is:

$T = \lambda$.

2. **The M/M/1/K queue**.

As in real life queues are always finite, we will further introduce the M/M1/K queue, Which can accommodate up to K customers, including the one in the service area, if the case.

Let's introduce the following notations:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- λ and μ are the rate of the Poisson process for arrivals and service times, respectively.

- $\pi(i)$, where $i=0,1,\ldots,K$, is the density function of the queue length in steady state.

In this case, the balance equations for this birth and death process are:

$\lambda \pi(0) = \mu \pi(1)$

$(\lambda + \mu)\pi(i) = \lambda \pi(i-1) + \mu \pi(i+1)$, for $i=1,2,\ldots, K-1$

$\lambda \pi(K-1) = \mu \pi(K)$.

Further, let's consider the notation

$$\rho = \frac{\lambda}{\mu}.$$

Then we have the following remarkable result for this type of queue:

$$\pi(i) = \frac{(1-\rho)\rho^i}{1-\rho^{K+1}}, \text{ for } i > K.$$

If $\lambda = \mu$ (i.e. $\rho = 1$), then:

$$\pi(i) = \frac{1}{K+1} \quad \text{for } i=0,1,\ldots,K$$

and

$\pi(i) = 0$ for $i > K$.

Therefore, the probability for an incoming customer to be rejected is $\pi(K)$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. The M/M/c queue

The particularity of this type of queue is that it has $c > 1$ servers and the waiting queue can be infinitely large.

Let's note λ and μ the rate of the Poisson process for arrivals and service times, respectively.

The number of customers in the system can be modeled as a birth and death process.

The birth rate is:

$\lambda_i = \lambda$, for $i \geq 0$.

The death rate is:

$\mu_i = i\,\mu$ for $i = 1,2,\ldots,c-1$ and

$\mu_i = c\,\mu$ for $i \geq c$

The density function modeling the queue length for this type of queue, is:

$\pi(i) = \pi(0)\,\rho^i / i!$ for $i = 0,1,\ldots,c$ and

$\pi(i) = \pi(0)\,\rho^i\, c^{c-i} / c!$ for $i \geq c$,

where:

$$\pi(0) = \frac{1}{\sum_{i=0}^{c-1} \frac{\rho^i}{i!} + \frac{\rho^c}{c!} \cdot \frac{1}{1-\rho/c}}.$$

The probability for a customer to stay in line (not be served immediately) is:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$P_{queuing} = \sum_{i=c}^{\infty} \pi(i) = \sum_{i=c}^{\infty} \pi(0) \frac{\rho^i c^{c-i}}{c!}$$ or, using the formula above:

$$P_{queuing} = \frac{(\rho^c / c!)(1/1-\rho/c)}{[\sum \rho^i / i! + (\rho^c / c!)(1/1-\rho/c)]}.$$

This formula is largely known in telecommunication under the name of **Erlang's C formula**, and expresses the probability that an incoming customer to not find a free trunk available, in a system of c trunks.

4. The M/M/c/c queue

This type of queue describes the situation in which we have $c > 1$ simultaneous servers but there is no waiting queue. Each new customer gets a private server but after c customers any other customer is lost (no waiting room).

The meaning of the parameters λ and μ are as above.

The number of the busy servers can be modeled as a birth and death process.

The birth rate is:

$\lambda_i = \lambda$ for $i = 0,1,\ldots,c-1$ and

$\lambda_i = 0$ for $i \geq c$.

The death rate is:

$\mu_i = i\,\mu$ for $i = 1,2,\ldots,c$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The density function of the number of busy servers is:

$$\pi(i) = \pi(0)\frac{\rho^i}{i!} \quad \text{for } i = 0,1,\ldots,c \quad \text{and}$$

$$\pi(i) = 0 \quad \text{for } i > c.$$

In this case

$$\pi(0) = \frac{1}{\sum_{i=0}^{c} \frac{\rho^i}{i!}}.$$

and

$$\pi(c) = \frac{\frac{\rho^i}{i!}}{\sum_{j=0}^{c} \frac{\rho^j}{j!}}.$$

This type of queuing system is also largely known in telecommunications as **Erlang's loss formula**, and it gives the probability that all trunks are busy.

Remarkable note: the formula for $\pi(i)$ above is true not only for exponential service time distribution but for any type of distribution. Therefore this property is called an insensitivity property.

5. The M/G/1 queue

In this type of queue:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- customers are served on the first-in, first-out (FIFO) rule.

- the arrival rate is still Poisson

- there is one server for the queue

- the serving time is a arbitrary probability density function G(x), the same for all customers.

This latest requirement means that, if s_i and s_j are the service times for two different customers, where $i \neq j$, then

- s_i and s_j are independent random variables, and

- $G(x) = P(s_i \leq x) = P(s_j \leq x)$, for all $x \geq 0$.

For this queuing system, the number of customers in the queue is not a Markov process. Let's note W_i to be the waiting time in the queue of the i-th customer, and we use the following notations:

W is the mean waiting time

$X(t)$ is the number of the customers in the waiting room at time t

$X(t_i)$ is the number of waiting customers just before the arrival of the i-th customer

$R(t)$ is the service time of the customer in the service at time t

t_n is the arrival time of the n-th customer for all $n \geq 1$

s_n is the service time for the n-th customer;

We note that: $E[s_n] = 1/\mu$.

We have:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$E[W_i] = E[R(t_i)] + \underset{j=i-X(ti)}{\overset{i-1}{E}}\left[\sum sj\right]$$

or:

$$E[W_i] = E[R(t_i)] + \frac{1}{\mu} E[X(t_i)] .$$

If $i \rightarrow \infty$ then we have:

$$W = R + X/\mu.$$

where:

$\underline{R} = \lim_{i \rightarrow \infty} E[R(t_i)]$ is the mean service time at arrivals, and

$\underline{X} = \lim_{i \rightarrow \infty} E[X(t_i)]$ is the mean number of the customers in the waiting room at arrivals.

If we apply Little's formula to the waiting room, for this type of queues, we have:

$$\underline{X} = \lambda \underline{W}$$

$$\underline{W}(1-\rho) = \underline{R},$$

so:

$$\underline{W} = \frac{\underline{R}}{1-\rho} .$$

Also one can prove that the mean waiting time in this type of queue is:

$$\underline{W} = \frac{\lambda E[s^2]}{2(1-\rho)}$$

where $E[s^2] = E[s_i^2]$ for all $i \geq 1$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and is named the **Polaczek-Khinchin formula**.

The mean system response time is given by:

$$\underline{T} = \frac{1}{\mu.} + \frac{\lambda E[s^2]}{2(1-\rho)}$$

The mean number of the customers in the system (waiting or being served) is:

$$\underline{N} = \rho + \frac{\lambda^2 E[s^2]}{2(1-\rho)}$$

Priority Queues

In some applications it becomes important to give some customers a higher priority in being served.

In a priority queuing system customers are divided in $K \geq 2$ classes, in which the lower the priority, the higher the class number. Usually it is assumed that customer within a class are served on a FIFO basis.

There two categories of priority policies:

- preemptive-resume policies, in which a service is interrupted whenever an arriving customer has a higher priority than the customer being served; the customer who's been interrupted, returns to the top of its priority queue and resumes service only if no customer of higher priority is in the waiting room.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(A variation of this policy is the preemptive-repeat policy in which an interrupted customer will have to restart service from the beginning).

- non-preemptive priority policies, in which a customer being served will be served regardless its level of priority and only after exiting the system, the servicing of the next customer will be prioritized.

1. Non-Preemptive Priority Queues

Let's consider a non-preemptive priority queue of type M/G/1 with K priority classes, in which customers of class k = 1,2,...,K arrive according to a Poisson process with rate λ_k and have independent service times with common density function $G_k(x)$ of mean $1/\mu_k$. Also, we assume that arrival time and service times are independent from each-other.

We introduce the following notations:

- $\rho_k = \lambda_k/\mu_k$ is the traffic intensity of customers from class k, k = 1, 2, ...,K.
- W_k is the waiting time, without being served, of a customer of class k
- X_k is the number of customers of class k, waiting in queue
- R is the residual service time of the customer in the server.

We have the following results:

$\underline{T}_k = \underline{W}_k + 1/\mu_k$ for k = 1; 2; : : : ;K, so we need only compute W_k.

for the highest priority customers we have:

$$\underline{W}_1 = \frac{\underline{R}}{1 - \mu_1} \quad \text{if } \rho_1 < 1.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\underline{W}_2 = \underline{R} + (1/\mu_1)\,\underline{X}_1 + (1/\mu_2)\,\underline{X}_2 + (1/\mu_1)\,\underline{Z}_1,$$

where $\underline{Z}_1$ is the mean value of the expectation time to serve all customers of class 1 that will arrive during the total waiting time of a class 2 customer.

From Little's formula we have in this case:

$$\underline{X}_1 = \lambda_1 \underline{W}_1$$

$$\underline{X}_2 = \lambda_1 \underline{W}_2$$

Other expressions of $\underline{W}_2$ are:

$$\underline{W}_2 = \frac{\underline{R} + \rho_1 \underline{W}_1}{1 - \rho_1 - \rho_2} \quad \text{or}$$

$$\underline{W}_2 = \frac{\underline{R}}{(1 - \rho_1)(1 - \rho_1 - \rho_2)} \quad \text{when } \rho_1 + \rho_2 < 1.$$

The general formula for $\underline{W}_k$ is:

$$\underline{W}_k = \frac{\underline{R}}{(1 - \sum_{j=1}^{k-1} \rho_j)(1 - \sum_{j=1}^{k} \rho_j)} \quad \text{when } \sum \rho_j < 1 .$$

We notice that $\sum_{j=1}^{K} \rho_j < 1$ represent the stability condition of this model.

At the same time, it can be proven that:

- the residual service time is:

$$\underline{R} = \sum_{j=1}^{K} \lambda_j \frac{\underline{\sigma}_j^{\,2}}{2}$$

where: $\underline{\sigma}_j^{\,2}$ is the mean service time for the customers of class j=1,..K.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- the mean waiting time of customers of class k is:

(for all k=1,2,...,K)

$$\underline{W_k} = \frac{\sum_{j=1}^{K} \lambda_j \underline{\sigma}_j^{\,2}}{2\left(1 - \sum_{j=1}^{k-1} \rho_j\right)\left(1 - \sum_{j=1}^{k} \rho_j\right)}$$

- the mean staying time of a customer of class k in the system is:

$\underline{T_k} = 1/\mu_k + \underline{W_k}$

- the mean number of customers of class k in the system is: (by using Little's formula):

$\underline{N_k} = \rho_k + \lambda_k \underline{W_k}$

2. Preemptive-Resume Priority

In a preemptive-resume policy a service is interrupted whenever an arriving customer has a higher priority than the customer being served; the customer who's been interrupted, returns to the top of its priority queue and resumes service only if no customer of higher priority is in the waiting room.

Let's keep all used notations from the previous section, and let's assume that

$\rho < 1$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

One can prove that:

- the mean staying time of a customer of class k in the system is:

$$\underline{T_k} = \frac{1}{1 - \sum_{j=1}^{k-1} \rho_j} \left(\frac{1}{\mu_k} + \frac{\sum_{j=1}^{k} \lambda_j \sigma_j^{\,2}}{2\left(1 - \sum_{j=1}^{k} \rho_j\right)} \right)$$

- the mean number of customers of class k in the system is: (by using Little's formula):

$$\underline{N_k} = \frac{1}{1 - \sum_{j=1}^{k-1} \rho_j} \left(\rho_k + \frac{\lambda_k \sum_{j=1}^{k} \lambda_j \sigma_j^{\,2}}{2\left(1 - \sum_{j=1}^{k} \rho_j\right)} \right)$$

Bibliography

A.1. Adan, Ivo, Jaques Resing. "Queuing Theory". Eindhoven University of Technology, The Netherlands. 2001

A.2. Cooper, Robert B. "Introduction to Queuing Theory". Second Edition. North Holland Press. 1981.

A.3. Lazowska Edward D., John Zahorjan, G. Scott Graham, Kenneth C. Sevcik. "Quantitative System Performance : Computer system analysis using queuing network models". Prentice-Hall, Englewood Cliffs, N.J. 1984.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A.4. Nain, Philippe. "Basic Elements of Queuing Theory: Applications to the Modeling of Computer Systems". Lecture notes. INRIA. France. 1998.

A.5. Vastola, Keneth S. "Performance Modeling and Analysis of Computer Communication Networks". Electrical Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, Troy, NY. At: http://networks.ecse.rpi.edu/~vastola/pslinks/perf/perf.html

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bibliography

References for Chapter 1

1.1. "Guglielmo Marconi", http://www.wsone.com/fecha/marconi.htm

1.2. "History of Wireless", http://www.wow-com.com/consumer/faq/articles.cfm?ID=101

1.3. http://www.kdischool.ac.kr/library/data/t00054.PDF

1.4. "Russia's Popov: Did he invent radio?", http://www.wsone.com/fecha/popov.htm

1.5. The official U.S. Government publication History of Communications - Electronics in the United States Navy (Bureau of Ships and Office of Naval History, 1963)

References for Chapter 2

2.1. Ahamed, Syed V., Victor B. Lawrence. "Intelligent Broadband Multimedia Network". Kluwer Academic Publishers. Boston, USA. 1997

2.2. Ahamed, Syed V., Victor B. Lawrence. "Design and Engineering of Intelligent Communications Systems". Kluwer Academic Publishers. Boston USA. 1997.

2.3. Balston, D., R. Macario. "Cellular Radio Systems". Artech House. Norwood, MA. 1994

2.4. Clarke, M. "Networks and telecommunications: Design and Operation". John Willey, New York, NY. 1997.

2.5. Garg, Vijay K., Joseph E. Wilkes. "Wireless and Personal Communications Systems". Prentice Hall. New York. 1996.

2.6. "Introduction to Wireless Systems". Wiley and Sons. 1999.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.7. Paetsch, Michael. "Mobile Communications in the US and Europe: Regulation, Technology, and Markets". Artech House, Boston, MA. 1993

2.8. Pahlavan, K. A. Levesque. "Wireless Information Networks". John Willey. New York, NY. 1995.

2.9. Rappaport, Theodore S. "Wireless Communications: Principles and Practice. Second Edition. Prentice Hall. 2001.

2.10. Stallings, William. "Data and Computer Communications". Prentice Hall. 2001

2.11. Webb, William. "The Complete Wireless Communications Professional". Artech House Publishers, Boston USA. 1999.

2.12. Young, Harry E. "Wireless Basics". Second Edition. Intertec. 1996.

References for Chapter 3

3.1. Black, Uyless. "Second Generation Mobile and Wireless Networks". Prentice Hall, Upper Saddle River, NJ. 1999

3.2. Paetsch, Michael. "Mobile Communications in the US and Europe: Regulation, Technology, and Markets". Artech House, Boston, MA. 1993

3.3 Papadimitriou, G. "Wireless Networks". Wiley and Sons. December 2002.

References for Chapter 4

4.1. Black, Uyless. "Second Generation Mobile and Wireless Networks". Prentice Hall, Upper Saddle River, NJ. 1999.

4.2. "Cellular Intersystem Operations (Rev. C)". Technical Report IS-41. EIA/TIA. 1995.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.3. Lin, Yi-Bing, Imrich Chlamtac. "Wireless and mobile Networks Architecture". John Willey and Sons, New York, NY. 2001

4.4 Papadimitriou, G. "Wireless Networks". Wiley and Sons. December 2002

Prasad, R. "CDMA for Wireless Personal Communications". Artech House. Norwood, MA. 1996.

4.5. Redl, S., M. Weber, M. Oliphant. "Introduction to GSM". Artech House. Norwood MA. 1995.

4.6 Snyder, Randall A. "Wireless Telecommunications Networking with ANSI-41". McGraw Hill. 2001.

References for Chapter 5

5.1. "IMT-2000 (G3G) Strategy and Key Technical Issues". The Next-Generation Wireless Communications Conference. Far Eastone, Taiwan. 1999.

5.2. "Introduction to cdma2000 Standards for Spread Spectrum Systems". http://www.3gpp2.org/Public_html/specs/C.S0001-0_v1.0.pdf

5.3. Lin, Yi-Bing, Imrich Chlamtac. "Wireless and mobile Networks Architecture". John Willey and Sons, New York, NY. 2001

5.4. Nokia Website:. www.nokia.com/networks/17/gprs//gprs1.html, 2000.

Harmonized Global 3G (G3G) Technical Framework for ITU IMT-2000 CDMA Proposal. May 1999.

5.5. Ojanpera, Tero, R. Prasad. "Wideband CDMA for 3rd Generation Mobile Communications". Artech House. 1998.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.6 Trillium Digital Systems. General Packet Radio Service (GPRS)

White Paper. May 1999.

5.7. UMTS official website. http://www.umtsworld.com/technology/fdd

5.8. Zvonar, Z., P. Jung. "GSM Communications Towards Third Generation Systems".

Kluwer. Berlin, Germany. 1998

References for Chapter 6

6.1 AT&T Labs. "4G Wireless Streaming Systems and Services"

http://www.research.att.com/news/2001/September/4G.html

6.2. http://www.s3.kth.se/radio/4GW/

6.3. http://www.s3.kth.se/radio/4GW/public/papers.html

6.4 Evans, B.G. Baughan, K. "Visions of 4G". http://www.ist-wsi.org/B_Evans.pdf.

6.5. http://www.fcc.gov/wrc03/files/docs/meeting/iwg/iwg_1/

wrc03_iwg_1_background_doc_13.pdf.

6.6. IEEE, Mohr, Werner. Becher, Reinhard. "Mobile Communications

Beyond Third Generation".

6.7. "IP for 3G". Wiley and Sons. 2003.

6.8 Karlson, B. "Wireless Foresight". Wiley and Sons. September 2003.

6.9. Mohr, W., W. Konhauser. "Access Network Evolution Beyond 3rd Generation Mobile

Communications". IEEE communications Magazine. 2000.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

References for Chapter 7

7.1. Ahamed, Syed V., Victor B. Lawrence. "Intelligent Broadband Multimedia Network". Kluwer Academic Publishers. Boston, USA. 1997

7.2. Ahamed, Syed V., Victor B. Lawrence. "Design and Engineering of Intelligent Communications Systems". Kluwer Academic Publishers. Boston USA. 1997.

7.3. Bellcore Special Report: "Advanced Intelligent Network release 1, Network and Operations Plan". SR-NPL-001623, Issue 1, June 1990.

7.4. Black, Uyless. "The Intelligent Network Customizing Telecommunication Networks and Services". Prentice Hall, Upper Saddle River, NJ. 1998.

7.5. Christensen, Gerry, P.G. Florack, R. Duncan, "Wireless Intelligent Networking", Artech House. 2001.

7.6. Faynberg, Igor, L. R. Gabuzda, M. P. Kaplan, N. J. Shah. "The Intelligent Networks Standards and Their Application to Services". McGraw-Hill. 1997.

7.7. Garrahan, J.J.. "Intelligent Network Overview", IEEE Communications Magazine 31(3), pg. 30-37. March 1993

7.8. IEEE Special Issue: "Marching Toward the Global Intelligent Network". IEEE Communications Magazine, 31(3). March 1993.

7.9. Lobley N.C. "Intelligent mobile networks". BT Technologies, J Volume 13, No2, April 1995.

7.10. Magendanz, Thomas, R. Popescu-Zeletin. "Intelligent Networks: Basic Technology, Standards and Evolution". Thomson Computer Press. 1996.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

References for Chapter 8

8.1. Asuncion, Santamaria, F.J. Lopez-Hernandez. "Wireless LAN: Standards and Applications". Artech House. 2001.

8.2 Gast, M. "802.11 Wireless Networks: The Definitive Guide". O'Reilly and Associates. 2002.

8.3 Geier, J. "Wireless LANs", Second Edition. SAMS. 2001.

8.4. IEEE 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification". 1997.

8.5. IEEE 802.11a "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High Speed Physical Layer Extension in the 5GHz band". 1999.

8.6. IEEE 802.11b "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High Speed Physical Layer Extension in the 2.4GHz band". 1999.

8.7. IEEEE. "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications". IEEE Standard 802.11, IEEE. June 1999.

8.8. http://www.wi-fiplanet.com/tutorials/article.php/2236611

8.9. http://www.hiperlan2.org.

8.10. Keller, R. "Wireless ATM for Broadband Multimedia Wireless Access: the ATM mobile project". IEEE Personal Communications, pg. 66-80. Vol.6 No.5, Oct. 1999.

References for Chapter 9

9.1. "Encapsulation within IP" - RFC-2003. IETF. Oct. 1996.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

9.2. http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/moblip/mobil_ip.pdf

9.3. Ioannidis, John, G.Q. Maguire Jr. "The Design and Implementation of a Mobile Internetworking Architecture". Proceedings of the Winter USENIX Conference". Jan. 1993.

9.4. "IPv4 Mobility Support" - RFC-2002. IETF. Oct. 1996.

9.5. Let, Hui, C. Perkins. "Ad-hoc networking with Mobile IP". Proceedings of the Second European Personal Mobile Communications Conference". Oct. 1997.

9.6. Perkins, Charles E. "Mobile IP, Design, Principles and Practices". Addison-Wesley. 1998.

References for Chapter 10

10.1. Ahamed, Syed V., Victor B. Lawrence. "Intelligent Broadband Multimedia Network". Kluwer Academic Publishers. Boston, USA. 1997

10.2. Ahamed, Syed V., Victor B. Lawrence. "Design and Engineering of Intelligent Communications Systems". Kluwer Academic Publishers. Boston USA. 1997.

10.3 Ganz, Aura, Zvi Ganz, Kitti Wongthavarawat. "Multimedia Wireless Networks: Technologies, Standards and QoS" Prentice Hall. 2003

10.4 Jha Sanjay, Mahbub Hassan. "Engineering Internet QoS". Artech House. 2002.

10.5. "Voice and Fax over IP". At http://www.iec.org/online/tutorials/vfoip/index.html

10.6. Wesel, Ellen K. "Wireless Multimedia Communications". Adison-Wesley wireless Communications Series. 1998.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

References for Chapter 11

11.1. Emmerson, Bob. David Greetham. " Computer Telephony and Wireless Technologies : Future Directions in Communications". Computer Technology Research Corporation. July 1997.

11.2. Minoli, D., E. Minoli. "Delivering Voice over IP Networks". John Willey and Sons. 1998.

11.3 Minoli, D. "Delivering Voice Over IP Networks". John Willey and Sons. Oct. 2002.

11.4 Miller, M. "Voice over IP Technologies". M.&.T PUBLISHING LTD. June 2002.

11.5. Rao, H., Y. Lin, S. Chou. "iGSM: VoIP Service for Mobile Networks". I.E.E.E. Communications Magazine, 4(38), pg. 62-69. 2000.

11.6. Reynolds, J. "Going Wi-Fi: A Practical Guide to Planning & Building an 802.11 Network". CMP Books. 2003.

References for Chapter 12

12.1. Fishwick, Paul. "Computer Simulation". At: http://www.cise.ufl.edu/~fishwick/introsim/paper.html.

12.2. Heidemann, John, N. Bulusu, J.Elson, et al. "Effects of Detail in Wireless Network Simulation". Distributed Systems Modeling and Simulation Conference. Jan. 2001.

12.3. Heidemann, John, K. Mills, S. Kumar. "Expanding Confidence in Network Simulations". IEEE Network. Sept/Oct. 2001.

12.4. http://ptolemy.eecs.berkeley.edu/ptolemyII/ptII3.0/ptII3.0/doc/

12.5. http://www.omnetpp.org

12.6. http://ptolemy.berkeley.edu/

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

12.7. http://www.isi.edu/nsnam/ns/

12.8. http://www.isi.edu/nsnam/ns/tutorial/

12.9. http://www.ssfnet.org/internetPage.html#ssfnet

12.10. http://www.scalable-networks.com/products/qualnet.php

12.11. http://deim.etse.urv.es/ajuda/manuals/comnet/c3ref.pdf

12.12. http://www.managementsoftware.hp.com/sso/isv/detail?appid=A668

12.13. http://jan.ucc.nau.edu/~ejacobs/Comnet.htm

12.14. http://www.opnet.com/products/modeler/home.html.

12.15. Johnson, D. B. "Validation of Wireless and Mobile Network Models and Simulation". Proceedings DARPA/NIST Network Simulation Validation Workshop. Fairfax, VA. May 1999.

12.16. Schwartz M. "Telecommunications Networks: Protocols, Modeling and Analysis". Addison-Wesley. Reading, MA. 1987

12.17. Zeigler, Bernard P., H. Praehofer, T.G. Kim. "Theory of Modeling and Simulation", Second Edition. Academic Press, San Diego. 2000.

References for Chapter 13

13.1. Leung, Kin K., B. McNair, L. Cimini, J.H. Winters. "Outdoor IEEE802.11 Cellular Networks: MAC Protocol Design and Performance". IEEE Communications Magazine. pg. 595-5999. Feb. 2002.

13.2. Lucas, Matthew T., B.J. Dempsey, D.E. Wrege, A.C. Weaver. "An Efficient background Traffic Model for Wide-Area Network Simulation". Proceedings of Sixth International Conference on Computer Communications and Networks. Sept. 1997.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

13.3. Motorola. Forrester, Nigel. "Measuring Performance and Quality of packet Voice Systems", Whitepaper. Jan. 2003.

13.4. Nacht Telecommunications Inc. "VoIP Industrial tutorial". April 11, 2001.

13.5. Ramakrishna, Sudhir, J. M. Holtzman. "A Scheme for Throughput Maximization in a Dual-Class CDMA System". IEEE Journal on Selected Areas in Communications, Vol. 16, No. 6, August 1998.

13.6. Sanneck, Henning. N.T. Long Le, M. Hardt, W. Mohr. "Selective Packet Prioritization for Wireless Voice over IP".

13.7. Veeraraghavan, N., N. Cocker, T. Moors. "Support of Voce Services in IEEE.802.11". Proceedings IEEE INFOCOM, Anchorage, AK. April 2001.

13.8. Sanneck, H., N. Le, A. Wolisz, G. Carle. "Interflow Loss Recovery and control for VoIP". Proceedings ACM Multimedia, Ottawa, Canada. Sept. 2001.

References for Appendix A

A.1. Adan, Ivo, Jaques Resing. "Queuing Theory". Eindhoven University of Technology, The Netherlands. 2001

A.2. Cooper, Robert B. "Introduction to Queuing Theory". Second Edition. North Holland Press. 1981.

A.3. Lazowska Edward D., John Zahorjan, G. Scott Graham, Kenneth C. Sevcik. "Quantitative System Performance : Computer system analysis using queuing network models". Prentice-Hall, Englewood Cliffs, N.J. 1984.

A.4. Nain, Philippe. "Basic Elements of Queuing Theory: Applications to the Modeling of Computer Systems". Lecture notes. INRIA. France. 1998.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A.5. Vastola, Keneth S. "Performance Modeling and Analysis of Computer Communication Networks". Electrical Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, Troy, NY. At: http://networks.ecse.rpi.edu/~vastola/pslinks/perf/perf.html

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.