INTERNATIONAL CONFERENCE & WORKSHOP

ON

THIRD GENERATION (3G) GSM & MOBILE COMPUTING: AN EMERGING GROWTH ENGINE FOR NATIONAL DEVELOPMENT

Covenant University

3GSMC 2007®

Edited by

A.A. ATAYERO, PhD

V.O. MATTHEWS, PhD

C.K. AYO, PhD

Published by :

Department of Electrical & Information Engineering,
College of Science and Technology,
Covenant University,
Ota, Nigeria.

3GSMC 2007®

Editors:

A.A. Atayero
V.O. Matthews
C.K. Ayo

E-MAIL: conference@covenantuniversity.com

POSTAL ADDRESS: Covenant University
Canaan Land, Km.10 Idiroko Road,
PMB 1023, Ota, Ogun State, Nigeria

WEBSITES: www.covenantuniversity.com

ISBN: 978-37963-2-1

Cover Design by *A.A.Atayero*

THE EDITORS

- ***ADEREMI A. ATAYERO*** is a Senior Lecturer in the Department of Electrical and Information Engineering, College of Science and Technology, Covenant University.
- ***VICTOR O. MATTHEWS*** is a Senior Lecturer in the Department of Electrical and Information Engineering, College of Science and Technology, Covenant University.
- ***CHARLES K. AYO*** is a Senior Lecturer and acting Head, Department of Computer & Information Sciences, Covenant University.

PREFACE

A major objective of setting up Covenant University is to eradicate the oppression of the black race engendered by the twin evil of ignorance and poverty. In support of this most noble of tasks, the department of Electrical & Information Engineering in unison with her sister department of Computer and Information Sciences deemed it fit to stage a conference-cum-workshop of this nature, where cutting-edge knowledge in two of the most dynamic fields of technological advancement (telecommunications and mobile computing) would be disseminated to the benefit of the participants. With this objective in mind, the Conference Planning Committee sought and found erudite scholars as well as key industry players to participate in the worthy cause of bringing Gown to Town and clothing Town in Gown.

This is a collection of papers presented by participants at the International Conference and Workshop on 3G GSM & Mobile Computing: An Emerging Growth Engine for National Development. The conference was organised by Covenant University, Ota and the Nigerian Communications Commission (NCC) in collaboration with key industry players in the GSM arena. It held in the Centre for Learning Resources at Covenant University, Canaanland, Ota, Nigeria, from 29–31 January, 2007.

This book of proceedings has been divided into two sections for ease of reference. These are: Part I – Invited Tutorial Papers and Part II – Technical Conference Papers. The papers presented in this book were subjected to the process of double blind peer-review and were found worthy of presentation at this conference by the team of seasoned reviewers.

It is the earnest hope of the Conference Planning Committee that both the participants of the conference and readers of this book of proceedings will find it a useful scholarly resource for reference purposes.

Thank you.

A.A. Atayero, PhD

Chairman, Conference Publication Committee.

ACKNOWLEDGEMENTS

The conference planning committee expresses its deep appreciation to the chancellor and the Chairman Board of Regents of Covenant University Bishop, Dr. David O. Oyedepo and the following personalities and corporations:

- Professor Aize I. Obayan – Vice Chancellor, Covenant University
- Dcn. 'Yemi Nathaniel – Registrar, Covenant University
- Professor James Katende – Dean, College of Science and Technology
- Professor T.S. Ibiyemi – University of Ilorin, Nigeria
- Professor A. Anoprieko – University of Donesk, Ukraine
- Engr. Ernest Ndukwe – CEO Nigerian Communications Commission
- Dr. Sylvanus Ehikioya – Nigerian Communications Commission
- Dr. Frank Ibikunle – M-tel, Nigeria
- The Managements of:
    - Ericsson Nigeria Limited
    - Huawei Technologies
    - Alcatel-Lucent
    - E-transact Nigeria

Most importantly, the conference planning committee appreciates the Lord Almighty for His provisions and protection during the planning and execution of this project.

CONFERENCE PLANNING COMMITTEE.

V.O. Matthews	Chairman
C.K. Ayo	Member
A.A. Atayero	Member
M.L. Akinyemi	Member
E.O. Agboje	Member
S.N. John	Member
J.O. Daramola	Member
N.A. Ikhu-Omoregbe	Member
O.S. James	Member
V.E. Ekong	Secretary

TABLE OF CONTENT

PREFACE 3

PART I 8

INVITED PAPERS:

SIMULATION TOOLS AND SERVICES FOR MOBILE USERS: HISTORY, STATE-OF-THE-ART AND FUTURE ....
PROFESSOR A. ANOPRIJENKO, S. JOHN, H. AL ABABNEH 9

GSM ARCHITECTURE ....
IBIKUNLE FRANK, PHD 21

UNIVERSITY AND THE REVOLUTION IN COMMUNICATIONS AND COMPUTING: EXPERIENCE OF DONETSK NATIONAL TECHNICAL UNIVERSITY
PROFESSOR A. ANOPRIJENKO, E. BASHKOV, AND A. MINAEV 34

CONTROL OF THE GSM RADIO RESOURCES ....
IBIKUNLE FRANK, PHD 47

PART II 57

TECHNICAL PAPERS:

COMPRESSING HIGH-DIMENSIONAL DATA SPACES USING NON-DIFFERENTIAL AUGMENTED VECTOR QUANTIZATION ....
O.O. OLUGBARA AND A.A. ATAYERO 58

AN ARCHITECTURAL FRAMEWORK FOR COLLABORATION OF HETEROGENEOUS COMMUNICATION DEVICES USING WAP AND MOBILE DEVICE AUGMENTED (MDA) GATEWAY INTEGRATION ....
A.A. AZETA, M.O. OYELAMI AND J.O. DARAMOLA 71

A PROCESS FRAMEWORK FOR SUBSCRIBER MANAGEMENT AND RETENTION IN NIGERIAN TELECOMMUNICATION INDUSTRY ....
J.O.DARAMOLA, O.O. OLADIPUPO AND A.G. MUSA 79

MOBILE INTERNET CRYPTOGRAPHY ....
T. OLORUNFEMI, B.K. ALESE AND O. AKINSELE 88

MOBILE PHONE: THE PAST, THE PRESENT AND THE FUTURE ....
M.O. OYELAMI AND A.A. AZETA 96

ONTOLOGIES AND SEMANTIC WEB SERVICES FOR M-COMMERCE APPLICATIONS ....
E.E.WILLIAMS, A.E. IBOR AND U.G. INYANG 105

A MULTI-CHANNEL APPLICATION FRAMEWORK FOR CUSTOMER CARE SERVICE USING BEST-FIRST SEARCH TECHNIQUE ....
O.M. ADEYEYE AND A.A. ATAYERO 113

RELIABILITY ISSUES OF GSM SECURITY PROTOCOL ....
O.S. OGUNDELE, O.S. ADEWALE AND B.K. ALESE 124

WIRELESS AND MOBILE NETWORKS ....
A.F. THOMPSON, B.K. ALESE AND O.S. ADEWALE 131

DEVELOPMENT AND SIMULATION OF AN ALGORITHM FOR DIMESIONING UMTS AIR INTERFACE FOR COVERAGE AND CAPACITY PREDICTION.................
O. Olabiyi, K. Ayodele, O. Ilori and K. Kehinde, ... 142

ON THE POWER SPECTRAL DENSITY OF THE GSM SIGNALING SCHEME ....
V.O. MATTHEWS AND U.A.K. OKONKWO 154

IMPACT OF MOBILE AND WIRELESS TECHNOLOGY ON HEALTHCARE DELIVERY SERVICES
Z.O. OMOGBADEGUN AND C.K. AYO 164

THE PROSPECTS OF M-VOTING IMPLEMENTATION IN NIGERIA ....
U.O. EKONG AND C.K. AYO 172

IMPLEMENTATION OF A LOW POWER SENSOR USING THE 1N4148 SIGNAL DIODE ....
F.E. IDACHABA AND V.E. EKONG 180

IMPLEMENTING OPERATIONS SUPPORT SYSTEMS IN E-HEALTH BASED SYSTEMS..
N.A. IKHU-OMOREGBE, S.A. EHIKIOYA, C.K. AYO 191

BROWSING THROUGH THE FOURTH GENERATION MOBILE NETWORKS............
O. SHOEWU, O. BADEJO AND O.T. BARUWA AND D.W.S. ALAUSA 260

ANALYSIS OF CAPACITY LIMITATION IN NIGERIAN GSM NETWORKS AND THE EFFECTS ON SERVICE PROVIDERS AND SUBSCRIBERS ....
V.O. MATTHEWS AND ADETIBA E. 208

THE FUTURE OF MOBILE INDUSTRY ....
R.S. ADESUYI 218

3G (GSM) TERMINAL CONCEPTS FOR RURAL AND URBAN EMPOWERMENT......
O. E. AGBOJE 233

SEARCH BY AUTHOR

PROFESSOR A. ANOPRIJENKO, S. JOHN, H. AL ABABNEH 9

IBIKUNLE FRANK, PHD.. 21

PROFESSOR A. ANOPRIJENKO, E. BASHKOV, AND A. MINAEV 34

IBIKUNLE FRANK, PHD.. 47

O.O. OLUGBARA AND A.A. ATAYERO 58

A.A. AZETA, M.O. OYELAMI AND J.O. DARAMOLA 71

J.O.DARAMOLA, O.O. OLADIPUPO AND A.G. MUSA.. 79

T. OLORUNFEMI, B.K. ALESE AND O. AKINSELE... 88

M.O. OYELAMI AND A.A. AZETA 96

E.E.WILLIAMS, A.E. IBOR AND U.G. INYANG.. 105

O.M. ADEYEYE AND A.A. ATAYERO 113

O.S. OGUNDELE, O.S. ADEWALE AND B.K. ALESE.. 124

A.F. THOMPSON, B.K. ALESE AND O.S. ADEWALE 131

O. OLABIYI, K. AYODELE, O. ILORI AND K. KEHINDE 142

V.O. MATTHEWS AND U.A.K. OKONKWO 154

Z.O. OMOGBADEGUN AND C.K. AYO 164

U.O. EKONG AND C.K. AYO 172

F.E. IDACHABA AND V.E. EKONG 180

N.A. IKHU-OMOREGBE, S.A. EHIKIOYA, C.K. AYO 191

O. SHOEWU, O. BADEJO AND O.T. BARUWA1 AND D.W.S. ALAUSA2 200

V.O. MATTHEWS AND ADETIBA E. 208

R.S. ADESUYI 218

O. E. AGBOJE 233

PART I

INVITED PAPERS:

WORKSHOP

SIMULATION TOOLS AND SERVICES FOR MOBILE USERS: HISTORY, STATE-OF-THE-ART AND FUTURE

PROFESSOR A. ANOPRIJENKO, S. JOHN, H. AL ABABNEH

Donetsk National Technical University, UA-83000, Artema, 58, Ukraine
anoprien@cs.dgtu.donetsk.ua
http://donntu.edu.ua

1 VERY LONG HISTORY OF SIMULATION FOR MOBILE USERS

One of the main ideas of this paper is a hypothesis that the simulation play an essentially more significant role in a human history and culture than it is usually assumed. On some examples it can be demonstrated that modern computational simulation has ancient prototypes and some artefacts can be interpreted as special simulation tools and environments. As typical examples of ancient simulation tools the "life/world tree" on mammoth bone and megalithical "models of the world" are presented. These artefacts were interpreted earlier as calendars, observatories or "ancient computers". The proposed hypothesis considers the following interpretation as most exact and appropriate: "special computational simulation tools and environments with real-time functions (calendar) and real-world interface (observatory)".

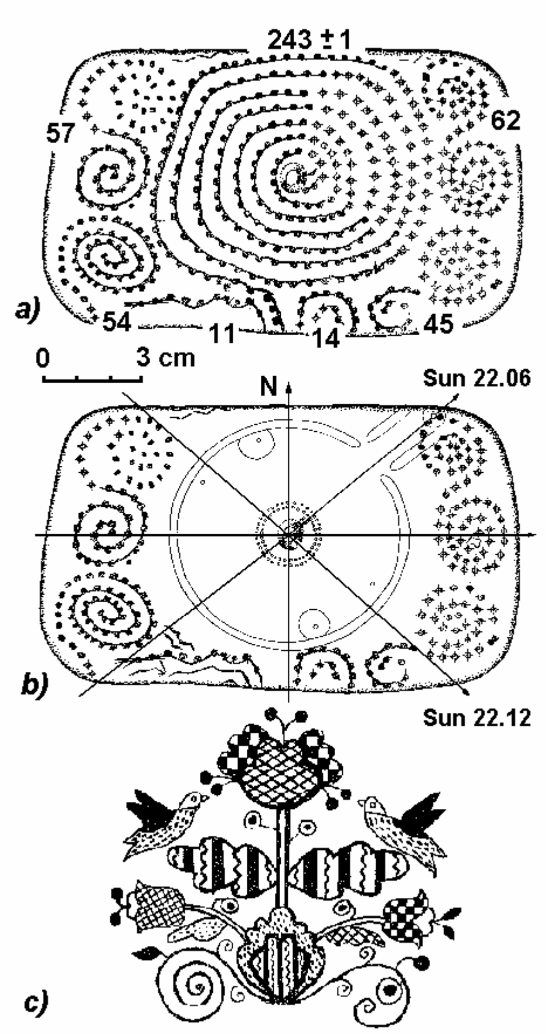


Figure 1. a) Quantitative characteristics of the "life/world tree" elements on the

We can say now about new science "archaeosimulation" [1-3]. The history of modern study of archaeosimulation begins from researches of Gerald S. Hawkins on Stonehenge more than 30 years ago (Hawkins and White 1966). Hawkins was not only the first who used a modern computer for the analysis of the ancient construction, but also he has declared the existence of "stone computers". Late other megalithic monuments were investigated and described, which were probably used as observatories and special original computers for registration and forecasting of the astronomical events (see, for example, Wood 1978). Almost all described structures can be interpreted as simulation tools.

As well as for modern science for ancient people the various forms of computational simulation were the most powerful means of research and understanding of complex dynamic processes of the real world.

New results described in the given work permit to interpret some other well known ancient artifacts as special tools for simulation.

In the Hermitage in St. Petersburg a small plate of mammoth bone with spiral figures of many dozens of dots is stored. It was found in 1929 in village Malta near the western part of the Baikal Lake (Siberia). The age of the plate is more than 15 thousand years.
While stored in the Hermitage the plate was periodically investigated by various scientists. One of the first was a German mythologist Karl Hentze. Hentze interprets spirals of the plate as symbols of the moon phases and even as an image of the whole universe, but without any quantitative analysis. The most careful analysis of semantics and quantitative system of the plate was done more than 10 years ago by Russian professor Larichev (Larichev 1989). His conclusions were as follows: advanced knowledge about the visible movements of the sky stars are fixed on the plate, which are a result of exact long-term observation of the Sun, Moon and visible planets. The precision of registration and representation of the information is quite enough for a sure prediction of the lunar and solar eclipse! Larichev has detected the following main elements on the plate:

- solar year: 243+62+45+14 = 365 days;
- lunar year: 243+57+54 = 354 days;
- four-year cycle: (242+63+45+14+11+54+58)x3 = 365.24 x 4 = 1461 days;
- sidereal form of the saros: 242x27,21=6585.35 days =18.61 solar years = 19 sidereal years;
- synodic form of the saros: (54+57+63+45+4)x29.53 = 6585.35 days;
- synodic cycle times for planets:

  Venus: (54+11+14+45) x 29.53 = 5 cycles;

  Mars: (62+57) x 29.53 = 4.5 cycles;

  Jupiter: (63+45) x 29.53 = 8 cycles;

  Saturn: (57+54+11) x 29.53 = 9.5 cycles.

Additional analysis of the plate as special simulation tool has allowed in determining the following:

1) The Malta plate model permits besides an exact “scientific” simulation of motions on the sky sphere also a simplified pragmatic “calendar” simulation for wide use:

- ≈1/6 of the solar year: 62 days;
- ≈1/8 of the solar year: 45 days;
- ≈ double sidereal month: 54 days;

- ≈ double sinodic month: 58 days;
- ≈ synodic cycle time for Mercury (four internal points of an element “14”): 4 x 29.5 = 116 days;
- ≈ synodic cycle time for Venus (ten external points of an element “14”): 10 x 29.5 x 2 = 590 days.

Then the plate can be interpreted as “model of the world” or “world tree”.

2) The element "14" can be easily used for observation of the female reproductive cycle:

Stage 1: 10 “external” days of barren period followed by menstruation.

Stage 2: (4+4) “internal” days followed by ovulation.

Stage 3: 10 “external” days before menstruation.

Stage 4: If menstruation does not come in time, then it will be necessary to make testy pass of the whole cycle (10+4+4+10).

Stage 5: In case of delay of the menstruation the cycle must be corrected.
Stage 6: If during the test pass of the cycle the menstruation was not, then go to central spiral “242”.

General term of pregnancy is 10+28+242=280 days.
Then central part of the plate can be interpreted as the “life tree”.

3) “Malta plate” was probably wide used special computational tool in ancient society, and it can be interpreted as specific simulational prototype for the famous mythological concept of “life/world tree” (Fig. 1c).

4) Baikal is located on the same latitude as Stonehenge. Main solar and lunar directions for Stonehenge and for the “mammoth plate” coincide. The plate could be used also as a "personal Stonehenge" or microobservatory (Fig. 1b).

5) Such form of fixing and transfer of the information allowed at the initial stage of history of the civilization (more than 10 000 years ago) to accumulate, apply and transmit knowledge without alphabet and another forms of writing.

Other good example of ancient simulation tool for mobile use is so called **Phaistos disk** (Fig.2). This is rectifying archaeological artifact from Crete. It is called the Phaistos Disc and is a small, baked clay disc about five inches in diameter. On the surface of each side of the disc there are incised spirals and within these are groups of hieroglyphic type characters which have never been interpreted [4].. The Phaistos Disc is thought to date to about

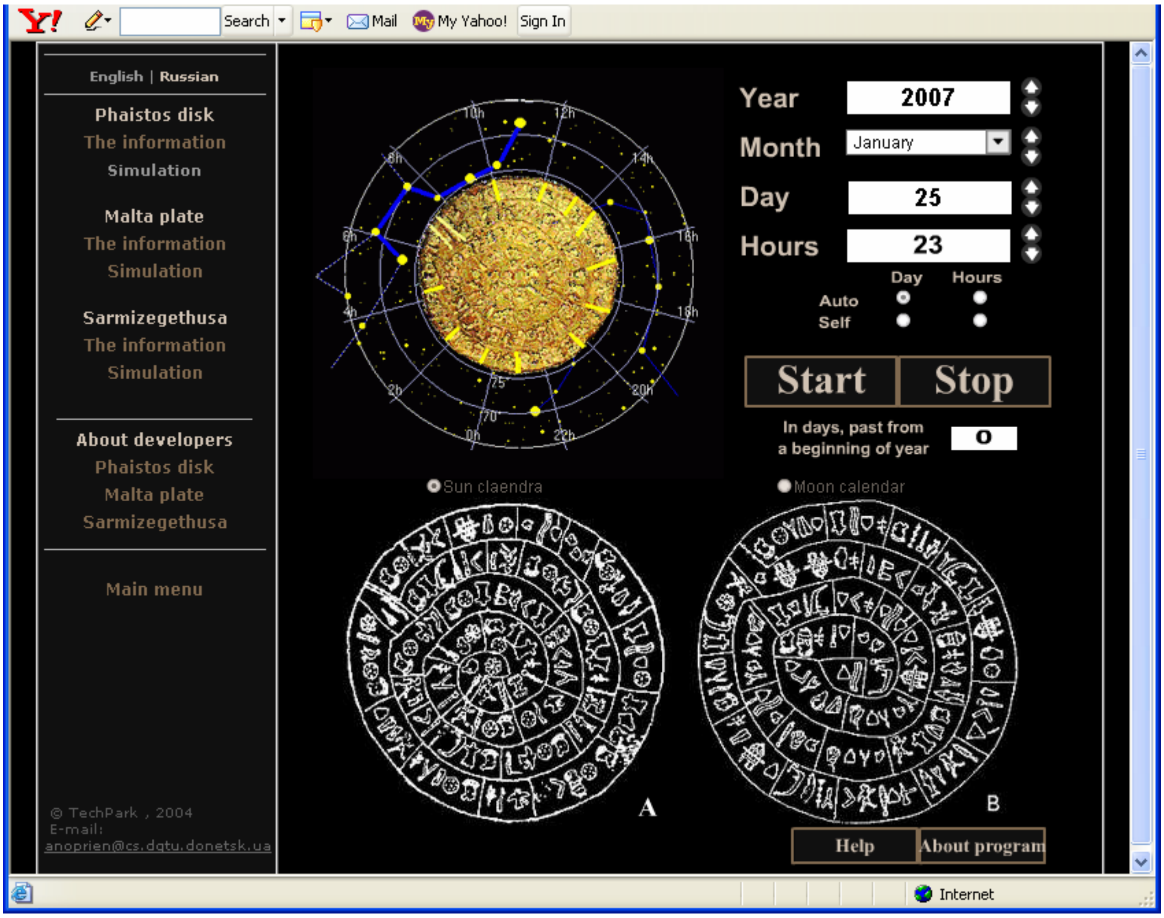


Figure 2. Simulation of Phaistos disk on on "Archeosimulation's portal"

1700BC. It is presently in the Heraklion Museum in Crete.

People have tried for decades to try and unravel the mystery of the Phasitos Disc but this has proved to be impossible for several reasons. There are very few examples of this sort of hieroglyphic script on Crete and since it never appears with any comparable text, there just are not enough examples for cross-referencing. Also there is no knowledge of the language spoken on Crete in Minoan times.

Each side of the Phaistos Disc contains an incised spiral. Within the whorls of these spirals are groups of hieroglyphs, pressed into the clay with pre-made tools. The groups are separated by incised lines crossing the spiral and there are anything between 2 and 8 glyphs per group. Side A of the Disc contains 123 glyphs within 31 groups, whilst side B has 119 glyphs within 30 groups. Many of the glyphs are repeated, some on both side A and side B of the Disc.

The Phaistos Disc and its groups of hieroglyphics was nothing more or less than a second calendar. This ran on periods of 123 days and was designed to allow the ritual year of 366 days to be corrected. Such a move is very important to any farming culture, so that planting and harvesting dates do not slip backwards of forwards throughout the year.

For both simulation artifacts we have new modern simulations tools which was developed in DonNTU (Ukraine) and now presented on “Archeosimulation’s portal” (http://simulation.in.ua/ asim1/index_en.htm). Such artifacts we can probably find in future also in Africa.

2 STATE-OF-THE-ART OF SIMULATION FOR MOBILE USERS

We can now use many various tools for simulation in mobile computers: Matlab/Simulink, GPSS, Excel, special tools for Computer Networks simulation (NetCracker and NS2 for example) and so on [5-8]. Most universal for mobile user now are various Java-tools. On the based of Java-platform we can develop simulation tool in wide range of applications: from mobile phones with J2ME to the global distributed environment with J2EE. Java-model for “Malta plate” was also developed in DonNTU.

A variant of expansion of classical binary logic and binary notation was proposed for simulation of all arts simulation tools from past, now and future: it is so called extended code-logical basis, which allow essentially increase information capacity and efficiency for description of dynamic processes and structures [9-10].

The majority of known modifications of binary and multivalued logic are in essence one-dimensional: the logic values 0 ("false", "no") and 1 ("true", "yes") are set on one axis, and all logic operations are carried out within the limits of this axis. But the variety of human knowledge is such, that the surrounding world can not be always unequivocally described by such a simplified scheme. Mainly because, there are at least two situations, which play an important part in the system of human knowledge: "Don't know" and "Know so much, that cannot answer definitely" [1]. An orthogonal arrangement of "False" and "True" axes permits to include the indicated situations in the formal logic system (Fig. 1). In such two-dimensional logic plane various systems of computer logic can be constructed.

The point A of absolute uncertainty is considered as origin of coordinates.

The points 1 and 0 are traditional logic significance’s “true” and “false”.

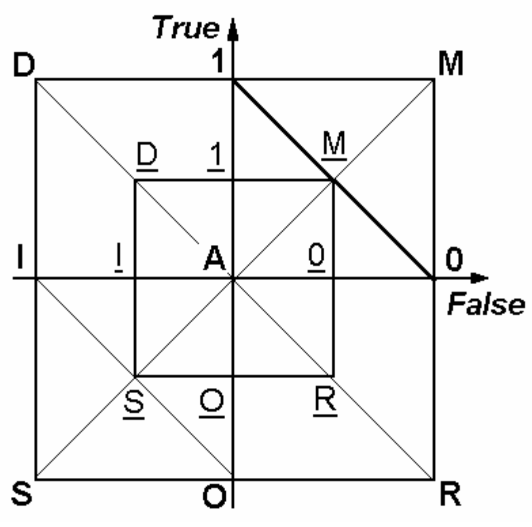


Figure 3. 2D logical space

M can be interpreted as “multivalue” or “true and false”.

$\underline{M}$ can be interpreted as “equiprobability of true and false”.

S - “symmetry” of M.

I and O - “symmetry” of 1 and 0 respectivly.
D and R - two others variants of the “multivalue”.

Various logical systems in this space we can specify as $L_N^K = \{x_1, x_2, ..., x_K\}$ with K= 1, 2, 3, 4, ... , where K is an order of the logic; $x_1, x_2, ..., x_K$ are various logical values from the 2D logical space; N - number of logic order K.

The variants of logic systems with four conditions present the heaviest theoretical and practical interest. Such systems can be named as "tetra logic". The tetra logic for example include except two traditional logic significance’s “true” and “false” also two additional significance’s:
$L_1^4 = \{1, 0, A, M\}$ (it was proposed in [1]), $L_2^4 = \{1, 0, \underline{M}, M\}$,

$L_3^4 = \{1, 0, S, M\}$, $L_4^4 = \{1, 0, A, \underline{M}\}$.

The introduction of the suggested logic conditions allows to essentially expand the opportunities of classical binary logic and to adapt it for features of human thinking.

3 THE ARCHITECTURE OF THE FUTURE DISTRIBUTED ENVIRONMENT

Intensive development of the Internet infrastructure and modern communications creates conditions for effective realization of various web-based simulation services. In our days there are two common ways of development such simulation services.
In the first case the system is develop from scratch, so it means that all the modern technologies can be used to gain better results. Of course it requires more time and efforts to development. Well known web-based simulation system built from scratch are: DEVSJAVA, JSIM, AnyLogic.

The other way is to use existent (legacy) simulation system to build fully functional web-based simulation environment. It means that we get all the benefits from the legacy system without necessity to develop it. The examples of such system are: NetSim, WSE, WebDIVA. Integration of simulation services into uniform simulation environments is the most perspective variant of their development. The architecture of such environment and its basic components are considered in the given article. The suggested architecture is generalization of long-term experience of models and environments that have been developed for simulation of complex dynamic systems [10-12]. One of the main goals of such system is to improve educational process [13]. Another one also, is the development of new web-based services on the basis of traditional simulation systems for a complex systems researching [14, 15].

Figure 4 shows the most general view future architecture of the Distributed Environment for Simulation Services Integration (DESSI) [16]. The basic elements of the environment are: Web-based Clients (WBC) which realizes the universal and specialized mobile user's interfaces for access to the integrated simulation environment. The basic technologies for realization of WBC are Java web-start clients, Java applets (for "normal" clients) and JavaScript/Flash/HTML for the so-called "lite" clients.

Simulation Services Provider (SSP) is the main "entrance points" for all remote queries from WBC. Basic function of SSP is keeping and granting the information about simulation services for WBC. Client can choose necessary service from the list of offered and initiate the beginning of simulation session.

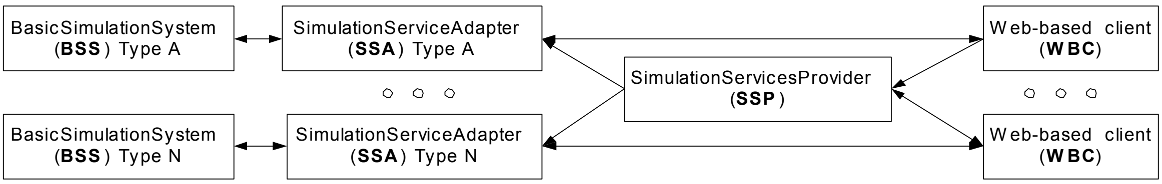


Figure 4. The overall view of the DESSI architecture

Simulation Service Adapter (SSA) is intended for communication with the base simulation systems. Every BSS normally has his own SSA. Basic function of SSA is providing the universal and transparent for the client access to the specific interfaces of various BSS. So the main role of SSA is implementing specific for each simulation system the mechanism of initialization, models loading, getting simulations results and so on.

The main benefits of such organizations are that for all the basic simulations systems (that were initially built using various languages and technologies) all clients have universal way to access simulations services.

Basic Simulation Environment (BSE) / Basic Simulation System (BSS) - are the main server components of the DESSI.

The UML-diagram of interaction between base modules is shown in figure 5 below.

To start the simulation session WBC client connects to the SSP provider. Based on the information received (for example, from descriptions in special XML file) SSP offers WBC accessible simulation services (Simulation Services, SS). Then user, according to his current tasks, chooses the most appropriate SS for him to make his work. Provider SSP creates a copy of SSA adapter and returns the reference of it to WBC. All further communications continue between WBC and SSA directly.

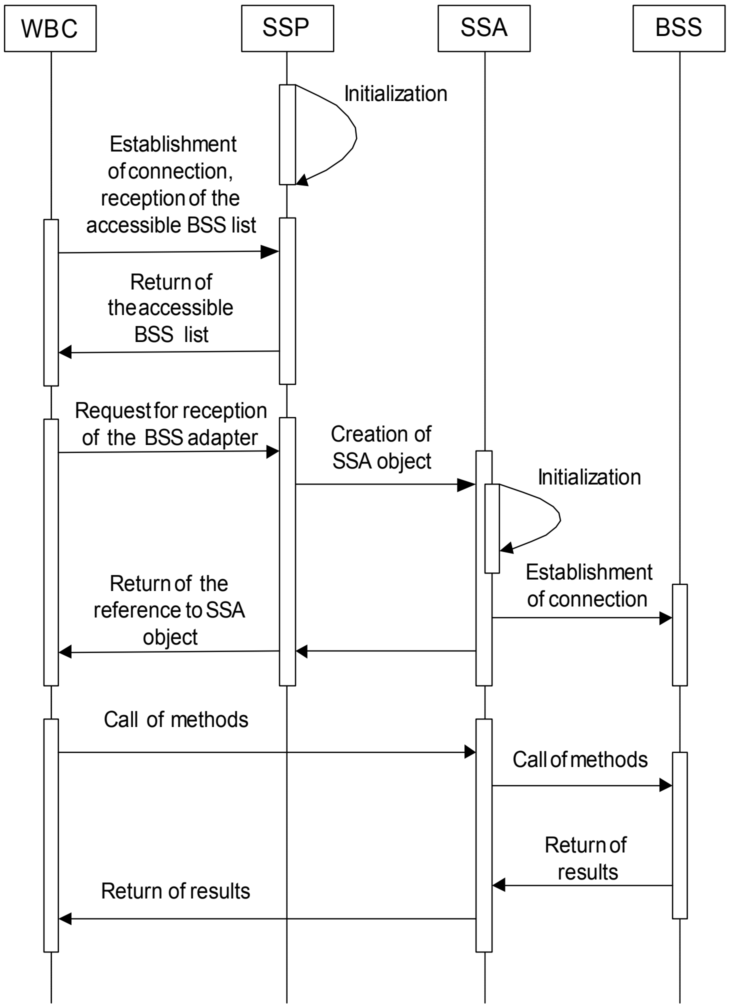


Figure 5. UML diagram of interaction between base modules of the DESSI

Simulation Service Factory (SSF). For the adapter (SSA) functioning, it is necessary to have stable and effective connection with a running copy of the base simulation environment, for example DIVA. For reception (and also clearing at the end of a session) this copy is meaningful to take advantage of pattern Factory. The basic idea is that Factory incurs all the operations connected to initialization and running a copy of the simulation environment, and also - covers the organization of objects pool that can give an essential gain in productivity (Fig. 6).

Common Object Request Broker Architecture (CORBA) is used for making communications with the BSS. It allows transparent inter-communications in object

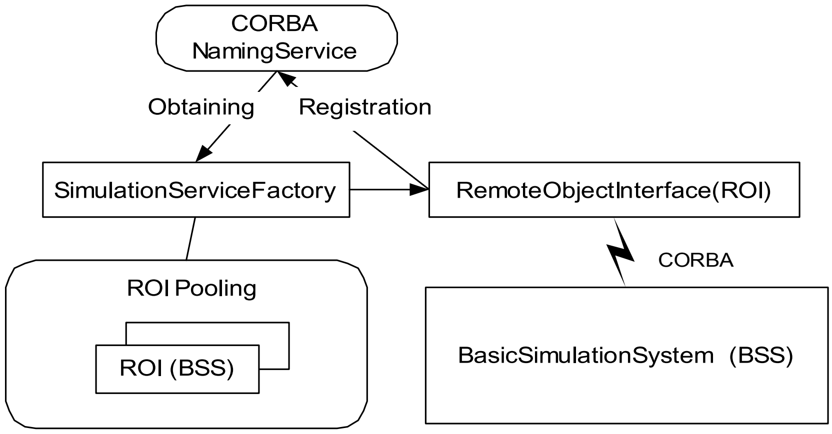


Figure 6. The organizations of objects "factory" for access to the simulation environment

oriented way between systems with various architecture types. A pool factory is formed from the remote object interfaces (Remote Object Interface, ROI).

The factory of object (SSF) is a special demon that works constantly during all life cycle of the system. The SSF does initialization and launching of several BSS copies and forms a pool of ROI objects in operative memory. The special XML-file is used to obtain the

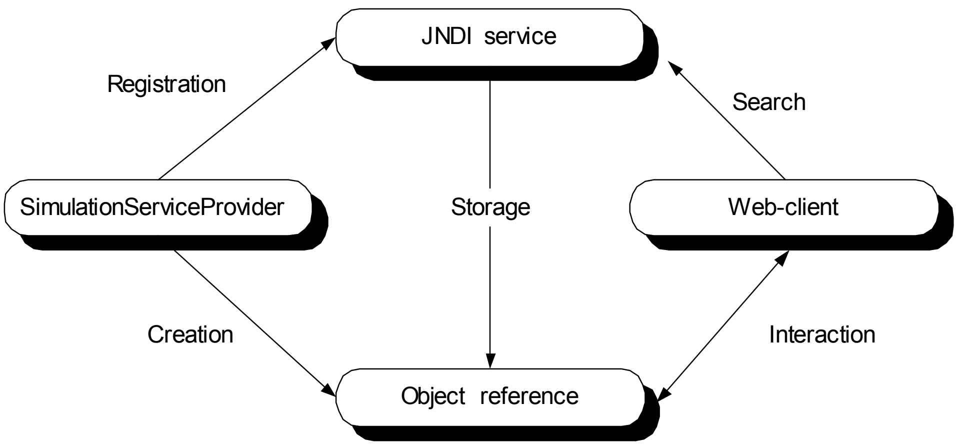


Figure 7. The mechanism of registration and search in JNDI service

necessary details to do it.

Registration of ROI objects in CORBA naming service (CORBA Naming Service) is necessary for further finding them by SSF (fig. 6).

The first and main point of entrance for the remote clients to the DESSI is Simulation Services Provider. It has theoretically infinite life cycle for processing client's requests. Traditionally, the server's demons implementation in "client-server" architecture is done by row TCP/IP sockets programming. It requires low-level and based on TCP protocol programming which for complex systems is not optimal.

Technology J2EE, namely specification JNDI, gives more flexible model for registration and search for the objects in the distributed systems (Fig. 7). The client is able to search for simulation services, using the information based only on a name of service, thus knowing the physical IP addresses of servers and numbers of TCP ports aren't necessary. During primary initialization SSP module registers itself in the JNDI service under name well-known for all DESSI clients. So it means that client will be allowed to receive the object references of SSP for the further interaction.

The main goal of DESSI is granting access for web-clients to the server-side simulation systems. The expansion of such systems will be the major and most a challenge in the next future. We already have good experience in development remote access and interaction with simulation system DIVA. Now we are in a process of developing a similar interaction with system Matlab/Simulink and others.

Figure 8 show the full model of the DESSI including all modules considered above.

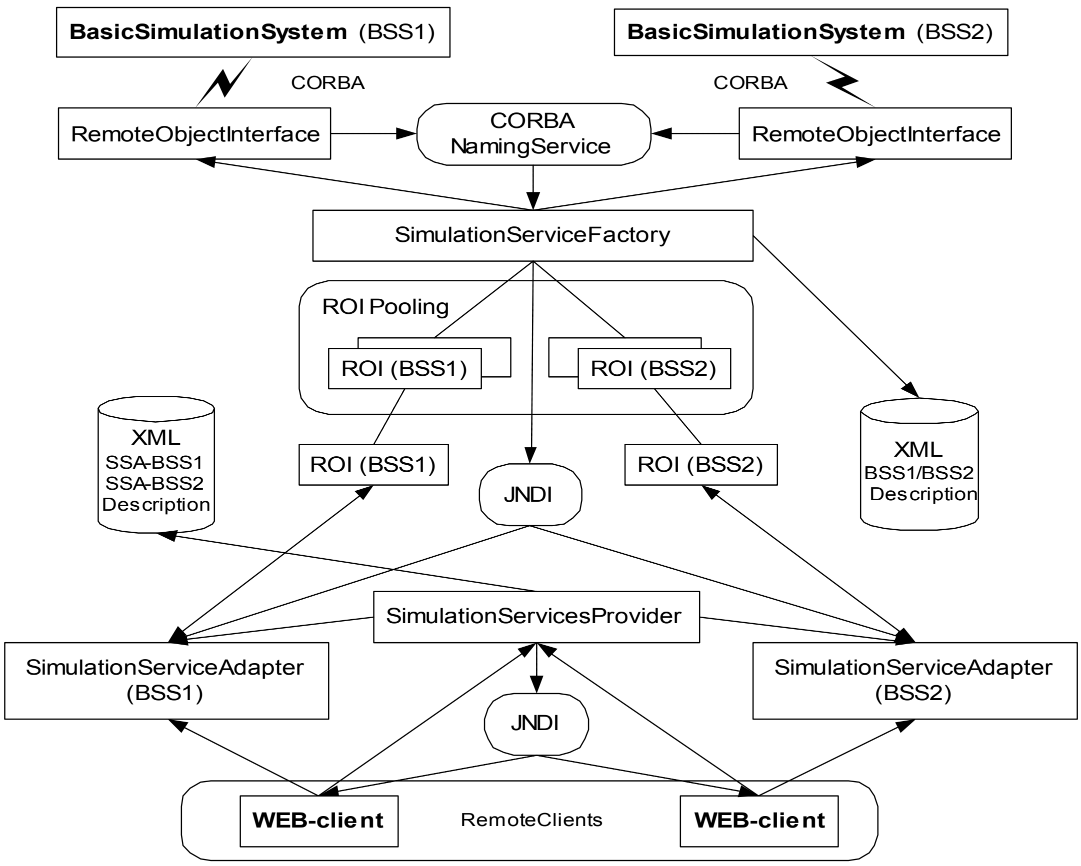


Figure 8. Structure of interaction between basic modules of DEISS (Example for 2 BSS)

The suggested architecture by now is the basis program in realization of the universal open WEB-based simulation environment with mobile client users. At first stage DESSI will give an opportunity for work with only complete models of various BSS. In supporting all the development cycles of models by BSE/BSS full control is planned as the further development of DESSI continues.

REFERENCES

1. Anoprienko A. Interpretation of some artifacts as special simulation tools and environments / “Short Papers Proceedings of the 1997 European Simulation Multiconference ESM'97. Istanbul, June 1-4, 1997" - Istanbul, SCS, 1997, p. 23-26.
2. Anoprienko A. Archaeosimulation: new sight on ancient society and lessons for computer era / Problems of Simulation and Computer Aided Design of Dynamic Systems. Scientific Papers of Donetsk State Technical University. Vol. 29. – Sevastopol: Weber, 2001. P. 320-326.
3. Anoprijenko A. The early history of simulation in Europe: scale planetariums and astromorphic models // EUROSIM 2004: 5th EUROSIM Congress on Modeling and Simulation. 06–10 September 2004. ESIEE Paris, Marne la Vallée, France. Book of abstracts. S. 146-147.
4. Butler, A. The Bronze Age Computer Disc // Foulsham & Co Ltd, March 1999
5. Waschler R., Kienle A., Anoprienko A., and Osipova T. Dynamic plantwide modelling, flowsheet simulation and nonlinear analysis of an industrial production plant // In J. Grievink and J. van Schijndel, editors, European Symposium on Computer Aided Process Engineering - 12 - ESCAPE-12, 26–29 May, 2002, The Hague, The Netherlands, Amsterdam: Elsevier, 2002, pages 583–588.
6. Waschler R., Kienle A., Sviatnyi S., Gilles E.D., Anoprienko A., Osipova T. Modeling and Simulation of a Chemical Reactor for the Production of Acetic Acid - III. Dynamic Phase Transitions // Problems of Simulation and Computer Aided Design of Dynamic Systems. Collected Volume of Scientific Papers. Donetsk State Technical University. Donetsk, 1999. - P. 102-109.
7. Svjatnyi V., Feldmann L., Lapko V., Anoprienko A., Reuter A., Bräunl T., Zeitz M. Massive parallel simulation of dynamic systems // Zeszyty naukowe. - 1997. - №1. - P. 207-229.
8. Anoprienko A., Svjatnyi V., Reuter A., Zeitz M. Simulation Systems on The Basis of a Parallel Computer / 8th European Simulation Symposium 1996 ESS'96, Genoa Italy, October 24th - 26th, 1996.
9. Anoprienko A. Tetralogic and tetracodes: an effective method for information coding // 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics. Berlin, August 24-29, 1997. Vol. 4. Artificial Intelligence and Computer Science. - Berlin: Wissenschaft und Technik Verlag. - 1997. - P. 751-754.
10. Anoprienko A., Svjatnyi V., Reuter A. Extended logical and numerical basis for computer simulation / “Short Papers Proceedings of the 1997 European Simulation Multiconference ESM'97. Istanbul, June 1-4, 1997" - Istanbul, SCS, 1997, p.
11. Svjatnyi V., Feldmann L., Lapko V., Anoprienko A., Reuter A., Bräunl T., Zeitz M. (1997) Massive parallel simulation of dynamic systems. Zeszyty naukowe, №1, 207-229.

12. Anoprienko A., Bazhenov L., Bräunl T. (1997) The development of the interface subsystem for the massive parallel simulation environment. 11. Symposium in Dortmund "Simulationstechnik" (pp. 672-677), ASIM-97, Vieweg, Braunschweig.
13. Minaev A., Bashkov E., Anoprienko A., Kargin A., Teslia V, Babasyuk A. (2002) Development of Internet Infrastructure for Higher Education in Donetsk Region of the Ukraine. ICEE 2002 Manchester International Conference on Engineering Education, Manchester, U.K.
14. Anoprienko A., Potapenko V. (2003) Web-basierte Simulationsumgebung mit DIVA-Serverkomponente für komplexe verfahrenstechnische Produktionsanlagen. 17. Symposium "Simulationstechnik" (pp. 205-208), ASIM 2003, Magdeburg.
15. Waschler R., Kienle A., Anoprienko A., and Osipova T. Dynamic plantwide modelling, flowsheet simulation and nonlinear analysis of an industrial production plant // In J. Grievink and J. van Schijndel, editors, European Symposium on Computer Aided Process Engineering - 12 - ESCAPE-12, 26–29 May, 2002, The Hague, The Netherlands, Amsterdam: Elsevier, 2002, pages 583–588.
16. Svjatnyj V., Anoprijenko A., Potapenko V., Zabrovsky S. The universal WEB-based distributed environment for simulation services integration // EUROSIM 2004: 5th EUROSIM Congress on Modeling and Simulation. 06–10 September 2004. ESIEE Paris, Marne la Vallée, France. Book of abstracts. S. 63-64.

GSM ARCHITECTURE

IBIKUNLE FRANK, PHD

M-TEL CHQ., Abuja, Nigeria

ABSTRACT

In late 80's emerged the second generation mobile systems. These systems were digital and mostly used Time Division Multiple Access (TDMA) techniques. Global System for Mobile Communications (GSM) was the most prominent amongst these and used a 13.6 kbps voice coding. Initially designed for 900 MHz operation, the systems are now available in 1800 MHz and 1900 MHz in the name of DCS 1800 or DCS 1900. The GSM system is by far the most dominant system used in the world.

1.0 INTRODUCTION

GSM, the Global System for Mobile Communications [1], [10] was developed to provide a global solution to the international need for global communications. It is the name of a standardization group established in 1982 to create a common European mobile telephone standard that would formulate specifications for a pan-European mobile cellular radio system operating at 900 MHz and 1800 MHz frequency bands. It uses Time Division Multiple Access (TDMA) techniques, where several different calls may share the same carrier. Each cell is assigned a particular time slot. Increase in demand and the poor quality of existing first generation cellular services led to ways to improve the quality of service and to support more users in the systems. Because the amount of frequency spectrum available for mobile cellular use was limited, efficient use of the required frequencies was needed for mobile cellular coverage. In GSM, rural and urban regions are divided into areas according to specific provisioning. Provisioning for each region is planned accordingly and includes cells, clusters, frequency reuse and handovers.

GSM, a digital cellular radio system involves dividing service area into regions called 'Cells'. A Cell is an area where a mobile station (MS) receives a signal strength that is high enough to set a radio connection on a dedicated channel (i.e. TCH) and maintain it. The size of a cell (or the size of area of coverage is mainly determine by the following parameters.

1. The output power at the antenna of BTS- Effective Radiated Power (ERP)
2. The minimum received level at the MS in busy mode
3. The minimum received level at BTS
4. The Timing Advance is the measure of the traveling time of the bursts between MS and BTS.

A cell must have exactly one BCCH that will carry essential information that must be know to the MS before call setup. The shape of a cell depend s on the antenna, which is connected to the cell. Maximum and minimum cell size is mainly determined by the distance radio waves can travel and the number of mobile subscribers who will be using the cell. Large cells are generally used in remote areas where there are very few subscribers (i.e. remote areas, high transmission power and few subscribers). The small cells are used in urban areas with low transmission power and many subscribers. A cell with its own cell sites is known as Omni-directional size. When traffic density is very high, cells can be reduced in size by sectorization (e.g. 120 degree and 60 degree sectors). This is achieved by focusing the antenna power on a certain sector of a circle. These sectored cells share the same cell site but each has its own allocation of radio carriers. To prevent interference between cells, a cell pattern called a cluster is designed. This pattern enable a frequency to be used only once with the aim of having a large frequency reuse distance. To enable allocation of channels to cells without co-channel interference, a re-use pattern is utilized.

The paper provides an introduction to GSM features, specifications, network architectures and an OSI model for the GSM system. A brief call sequences to and from a mobile station (MS) and a telephone line is also provided in order to set the context for understanding traffic cases in the GSM network.

2). KEY FEATURES OF GSM

GSM cellular telephone systems provide the MS subscriber and network providers with many advantages over a standard telephone network [7]. These features are as follows:
1) Fully Digital System: It incorporates error correction, thus protecting the traffic that it carries and uses radio resources much more efficiently than the analogue cellular systems.
2) Subscriber Identity Module
3) Complete system specifications and compatibility
4) Noise Robust: enable the use of tighter frequency re-use patterns and minimizing interference problem
5) Flexibility in providing international roaming; and Increase Capacity five (5) times that of analog cellular FDMA system by the use of a modulation technique called GMSK and a frequency reuse method
6) Use of standardized open interfaces (e.g., C7 and x.25): It uses standard CCITT protocols and interfaces that enable different pieces of hardware from different manufactures be compatible
7) Improved security and confidentiality: It offers enhanced privacy to subscribers and security to network providers
8) Provide cleaner and flexible handovers processes
9) Offers an enhanced range of Services: Speech Services; Data Services and Supplementary Services.

Table 1. GSM Air Interface Parameters and Specifications

SYSTEM	**GSM 900**
FREQUENCIES: UPLINKS DOWNLINKS	890 – 915 MHz 935 – 960 MHz
MODULATION TECHNIQUE	GMSK
MODULATION DEPTH	B.T = 0.3
MULTIPLE ACCESS	TDMA
DUPLEX	FDD
CHANNEL BANDWIDTH	25 MHz
Rx/Tx Spacing - Frequency	45 MHz
Rx/Tx Spacing – Time	1.15 msec (3 timeslots)
CHANNELS PER CARRIER (FR)	8
CARRIER SPACING	200 khz
SUBS CAP/RADIO Channels	125 to 300 Conversations per cell
CHANNEL DATA RATE	270 kbps
TDMA FRAME PERIOD	4.62 msec
TIMESLOT PERIOD	0.577msec
CODED CHANNEL RATE	0.58 msec
SPEECH CODEC	22.8 kbps
UNCODED VOICE RATE	RPE-LPT
BIT PERIOD	3.692us
BIT Per TIMESLOT	156.25 bits

3) THE GSM NETWORK COMPONENTS

The diagram in figure 1 shows a simplified GSM network. Each network component is designed to communicate over an interface specified by the GSM standards. The principle component groups of a GSM network are: the Mobile Station, Base Station System and Network Switching System.

3.1) Mobile Station (MS)

The MS consists of two parts, the Mobile Equipment (ME) and an electronic 'smart card' called a Subscriber Identity Module (SIM) that is plug into the ME and identifies the MS subscriber and also provides other information regarding the service that subscriber should receive.

The ME is the hardware used by the subscriber to access the network. The hardware has an identity number called the International Mobile Equipment Identity (IMEI) associated with it, which is unique for that particular device and permanently stored in it and enables the network operator to identify mobile equipment which may be causing problems on the system. There are three main types of ME:

Vehicle Mounted - These devices are mounted in a vehicle and the antenna is physically mounted on the outside of the vehicle. Portable Mobile Unit - This equipment can be handheld when in operation, but the antenna is not connected to the handset of the unit. Handheld Unit - This equipment comprises of a small telephone handset not much bigger than a calculator. The antenna is been connected to the handset. The ME is capable of operating at a certain maximum power output dependent on its class, type and use. These mobile types have distinct features which must be known by the network, for example their maximum transmission power and the services they support.

TABLE 2. MS POWER CLASSES

POWER CLASS	TYPE	MAX. OUTPUT POWER
1	Vehicle & Portable	Undefined
2	Vehicle & Portable	8 Watts
3	Handheld	5 Watts
4	Handheld	2 Watts
5	Handheld	0.8 Watts

The SIM as stated earlier is a "smart card" which plugs into the ME and contains information about the MS subscriber and contains several pieces of information such as: International Mobile Subscriber Identity (IMSI) - This number identifies the MS subscriber. It is only transmitted over the air during initialization. Temporary Mobile Subscriber Identity (TMSI) - This number identifies the subscriber, it is periodically changed by the system management to protect the subscriber from being identified by someone attempting to monitor the radio interface. Location Area Identity (LAI) - Identifies the current location of the subscriber. Subscriber Authentication Key (Ki) - This is used to authenticate the SIM card. Mobile Station International Services Digital Network (MSISDN) - This uniquely identifies a mobile telephone subscription in the PSTN numbering plan. This is the number dialed when calling a mobile subscriber. It is the actual telephone number of the subscriber and the only network identity that subscribers are aware of. Most of the data contained within the SIM is protected against reading (Ki) or alterations (IMSI). Others are:

Cell Global Identity (CGI) - This is used for identifying individual cells within a LA. Cell identification is achieved by adding a cell identity (CI) to the LAI components. The CI has a maximum length of 16 bits. Base Station Identity Code (BSIC) - The BSIC enables MSs to distinguish between different base stations sending on the same frequency. Location Number (LN) - This is a number related to a certain geographical area, which the network operator specifies by “tying” the location numbers to cells, location areas, or MSC/VLR service areas. It’s used to implement features like regional / local subscription and geographical differentiated charging.

Most of the data contained within the SIM is protected against reading (Ki) or alterations (IMSI). Some of the parameters (LAI) will be continuously updated to reflect the current location of the subscriber. The SIM card and the high degree of inbuilt system security, provide protection of the subscriber’s information and protection of networks against fraudulent access. SIM cards are designed to be difficult to duplicate. The SIM can be protected by use of Personal Identity Number (PIN) password, similar to bank/credit charge cards, to prevent unauthorized use of the card. The SIM is capable of storing additional information such as accumulated call charges. This information will be accessible to the customer via handset/keyboard key entry.

3.2) Base Station System (BSS)

The BSS provides the radio interconnection from the MS to the land-based switching equipment. It comprises a combination of digital and RF equipment. The BSS is responsible for channel allocation, link quality and power budget controls, signaling and broadcast traffic controls, Frequency Hopping and handover initiation. BSS communicates with the MS over the digital air interface (Um) and with the MSC via 2 Mbps links. BSS consists of three (3) major hardware components:
(i) The BSC has the digital control function of all BTSs. Between the BTS and BSC there is an access via the A-bis interface. Any operational information required by the BTS will be received via the BSC. Likewise any information required about the BTS (by the OMC for example) will be obtained by the BSC. The BSC incorporates a digital switching matrix, which it uses to connect the radio channels on the air interface with the terrestrial circuits from the MSC. The BSC switching matrix also allows the BSC to perform “handovers” between radio channels on BTSs, under its control, without involving the MSC.
(ii) The BTS is controlled by a BSC and contains one or more transceivers and provides the air interface connection with the MS and the network, by providing radio coverage functions from their antenna. (iii) The Transcoder is used to compact the signals from the MS so that they are more efficiently sent over the terrestrial interfaces, or in other word, reduces the rate at which the traffic (voice/data) is transmitted over the air interface. It is required to convert the speech or data output from the MSC (64 kbit/s PCM), into the form specified by GSM specifications for transmission over the air interface, that is, between the BSS and MS (64 kbps to 16 kbps and vice versa).

3.3) Network Switching System (NSS)

The NSS consists of the Mobile Switching Centre that performs the main switching functions of the GSM network, the databases required for subscriber data and mobility management, and processors together with the required interfaces. The components of the NSS are: Mobile Services Switching Centre; Home Location Register; Visitor Location Register; Equipment Identity Register; Authentication Centre; Message Center; Mobile Intelligent Network, Operation and Maintenance Centre, etc.
Mobile Switching Centre (MSC) - The MSC represents the gateway to other networks such as the Public Switched Telephone Network (PSTN), Integrated Services Digital Network (ISDN). The MSC functions include: paging, MS location updating, handover control. The MSC is accessible via four interfaces. The B/C/D interfaces are between the

MSC and the data location registers, and finally, the E interface is towards a neighboring MSC.
Gateway Mobile Switching Centre (GMSC) - Any MSC in the mobile network can function as a gateway by integration of the appropriate software and definition of HLR interrogation information. GMSC finds and interrogates HLR for roaming number and route the call according to the interrogation.
Home Location Register (HLR) - The HLR is the reference database for subscriber parameters. Various identification numbers and addresses are stored, as well as authentication parameters. The HLR database contains the master database of all the subscribers to a GSM PLMN. The data it contains is remotely accessed by all the MSCs and the VLRs in the network. The data can also be accessed by an MSC/ VLR in a different PLMN, to allow inter-system and inter-country roaming.
Visitor Location Register (VLR) - The VLR contains a copy of most of the data stored at the HLR. It is, however, temporary data which exists for only as long as the subscriber is "active" in the particular area covered by the VLR. The VLR database will therefore contain some duplicate data as well as more precise data relevant to the subscriber remaining within the VLR coverage. The additional data stored in the VLR are mobile status (busy/free/no answer etc.); Location Area Identity (LAI); Temporary Mobile Subscriber Identity and Mobile Station Roaming Number.
Equipment Identity Register (EIR) - The EIR contains a centralized database for validating the International Mobile Equipment Identity (IMEI). This database is concerned solely with MS equipment and not with the subscriber who is using it to make or receive a call. The EIR database consists of lists of IMEIs (or ranges of IMEIs) organized as follows: White List: contains those IMEIs which are known to have been assigned to valid MS equipment; Black List: contains IMEIs of MS which have been reported stolen or which are to be denied service for some other reason; Grey List: contains IMEIs of MS which have problems (for example, faulty software). The EIR database can be accessed by the MSCs in the network and or, by an MSC in a different PLMN.
Authentication Centre (AUC) – The AuC is co-located with the Home Location Register as it will be required to continuously access and update, as necessary, the system subscriber records. The authentication process will usually take place each time the subscriber "initializes" on the system.
Interworking Function (IWF) - The IWF provides the function to enable the GSM system to interface with the various forms of public and private data networks currently available. The basic features of the IWF are data rate adaptation and Protocol conversion. The IWF also incorporates a ''modem bank" which may be used when, for example, the GSM Data Terminal Equipment (DTE) exchanges data with a land DTE connected via an analogue modem.
Message Center (MC) - A Message Center (MC) may be added to a GSM network to provide voice/fax mail and handling of the Short Message Service (SMS) and SMS Cell Broadcast (SMSCB) text messages. These services can generate considerable revenue for a network operator, as they are becoming increasingly popular and controlled by OSS.
Mobile Intelligent Network (MIN) nodes – The MIN nodes when added to a basic GSM network provides value-added services to subscribers. The MIN nodes include:
- Service Switching Point (SSP): acts as an interface between the call control functions of the mobile network and the service control functions of a Service Control Point (SCP). SSP may be integrated within an MSC/VLR or as a stand-alone.

- Service Control Point (SCP): contains the intelligence of a MIN service or services. This intelligence is realized in software programs and data. The SCP is recommended to be a stand-alone node, accessible by all MSC/SSPs.
- Service Data Point (SDP): manages the data which is used by an MIN service. SDP is a stand-alone node based on UNIX.

3.4) Operations and Maintenance System

The operations and maintenance system provides the capability to manage the GSM network remotely. This area of the GSM network is not currently tightly specified by the GSM specifications, it is left to the network provider to decide what capabilities they wish it to have. The Operations and Maintenance System comprises of two parts:

Network Management Centre (NMC) - The Network Management Centre (NMC) has a view of the entire PLMN and is responsible for the management of the network as a whole. It resides at the top of the hierarchy and provides global network management. The NMC offers the ability to provide hierarchical regionalized network management of a complete GSM system. It is responsible for operations and maintenance at the network level, supported by the OMCs which are responsible for regional network management. The NMC is therefore a single logical facility at the top of the network management hierarchy with a high level view of the network. The NMC can take regional responsibility when an OMC is not manned, with the OMC acting as a transit point between the NMC and the network equipment.

Operations and Maintenance Centre (OMC) - is a centralized facility that supports the day to day management of a cellular network as well as providing a database for long term network engineering and planning tools. It filter information from the network equipment for forwarding to the NMC, thus allowing it to focus on issues requiring national co-ordination. An OMC manages a certain area of the PLMN thus giving regionalized network management. There are two types of OMC: OMC(R)- this controls specifically the Base Station System; and OMC(S) - this controls specifically the Network Switching System. The OMC supports the Event management, Alarm Management, Fault Management, Performance Management, Configuration Management and Security Management.

Billing GateWay (BGW)

A BGW collects billing information or Call Data Record (CDR) files from from the network elements and immediately forwards these to post-processing systems that use CDR files as input. The Billing Gateway (BGW) acts as a billing interface to all network elements in a GSM network.

4) GSM INTERFACES AND SIGNALING PROTOCOLS

4.1) GSM Terrestrial Interfaces

The terrestrial interfaces comprise all the connections between the GSM system entities apart from the air interface (Um). The terrestrial interface which acts as the physical bearer transport the traffic across the system and allow the passage of the thousands of data messages necessary to make the system function. They transport the data for software downloads and uploads the collection of statistical information and the implementation of operations and maintenance command. The four (4) standard terrestrial interfaces used are:

4.1.1) 2 Mbps - This carries traffic from the PSTN to the MSC, between MSCs, from an MSC to a BSC and from a BSC to a remotely sited BTSs. Each 2.048 Mbps link provides

thirty (30) 64 kbps channels available to carry speech, data or control information. The control information may contain C7, LAPD or X.25 formatted information.
4.1.2) Signaling System #7 (C7 or SS7) - This interface is responsible for carrying signaling and control information between most major entities, and to and from the PSTN.
4.1.3) X.25 Interfaces (Packet Switched Data) - The X.25 packets provide the OMC with communications to all the entities over which it has control and oversight. The X.25 connection is contained within 2 Mbps links using a dedicated timeslot. X.25 connection from the OMC to the BSS may be connected by software at the MSC.
4.1.4) A-bis - Because of the specific nature of the signaling and control information passing over the 2 Mbps links between the BSC and remotely sited BTS, a different type of interface is required. The GSM specification for this interface (termed Abis) is not very specific and therefore interpretations of the interface vary. This means that one manufacturer's BTS will not work with another manufacturer's BSC.
4.2) GSM Signaling Protocols

The signaling protocols used between GSM networks are:
(i) The X.25, (LAPB) Protocol - used between the BSC-OMC. (ii) The SS7 link, 1x 64 kbps timeslot (BSSAP, MAP, TCAP, SCCP, MTP) - used between the BSC - MSC, dependent on what type of signaling is required will depend on which part of the C7 Protocol will be used (for example, MSC-MS will be a subset of BSSAP called DTAP to Transfer Messages). (iii) The LAPD protocol, 1 x 64 kbps timeslot – is used between the BSC- BTS, this is normally 64 kbps but some manufacturers offer 16 kbps links as well. The link between the BSC-CBC does not use a specific protocol. The choice of protocol is decided between the PLMN and the CBC providers (typically X.25 or C7 may be used).

5) THE OSI REFERENCE MODEL

A pattern for structuring data communication networks in general has been provided by the International Standards Organization (ISO) in the form of the Open Systems Interconnection (OSI) model [11]. The OSI model provides for a number of horizontal layers, each layer communicating exclusively, and according to well defined rules, with the layers immediately above and below it. Communications thus becomes vertical, rather than horizontal, except for the lowest, or physical layer, where the information is passed from one system to the other. The GSM specifications or recommendation fully defined the lowest three layers of this OSI model.

In the lowest layer, layer 1, the physical characteristics of the transmission or radio path medium are specified, In the context of GSM radio link, this definition includes not only the frequencies, modulation type, etc., but also the structure of the bursts and frames implicit in a time division multiplex transmission scheme. This layer is responsible for the correct transmission of single bits, an element of error- protection coding also belongs here.

TABLE 3: OSI model and the use of three lowest layers (1 to 3) in GSM

OSI LAYER NO.	OSI LAYER NAME	GSM EQUIVALENT MODEL	TASKS

7	Application	-	User tasks
6	Presentation	-	
5	Session	-	Network Tasks
4	Transport	-	
3	Network	-Call management -Mobility management -Radio resource	
2	Data link	-Concatenation -Segmentation -Acknowledgement	GSM Network tasks
1	Physical	-Error detection -Channel coding -Modulation	

Layer 2, or the data-link layer, consists of an intelligent entity responsible for the safe communication of messages or frames between radio stations. To this end the transmitting section structures the messages of the higher layer to match the physical constraints of the layer 1 medium and request, in many situations, a confirmation (acknowledgement) from the receiving end. At the receiving end messages are reconstructed from the received frames and the acknowledgement are formulated for retransmission.

Layer 3, the network layer, is responsible for management of all calling and related activity of the radio network. These tasks are further subdivided into sub-layers designated call control management, mobility management and radio resource management.

The higher layers apply to any telecommunications system and specific reference to these tasks does not appear in the GSM recommendations.

6) GSM BASIC CALL SEQUENCES

This section and figure 3 describe the basic components and processes involved in setting up a call between a GSM mobile and a subscriber in the PSTN, and from PSTN to a mobile subscriber [4].

6.1) CALL FROM AN MS PROCEDURES

1. The MS uses RACH to ask for a signaling channel.
2. The BSC/TRC allocates a signaling channel, using AGCH.
3. The MS sends a call set-up request via SDCCH to the MSC/VLR. Over SDCCH all signaling preceding a call takes place. This includes:
   - Marking the MS as "active" in the VLR
   - The authentication procedure
   - Start ciphering
   - Equipment identification
   - Sending the B-subscriber's number to the network
   - Checking if the subscriber has the service "Barring of outgoing calls" activated
4. The MSC/VLR instructs the BSC/TRC to allocate an idle TCH. The BTS and MS are told to tune to the TCH.
5. The MSC/VLR forwards the B–number to an exchange in the PSTN, which establishes a connection to the subscriber.
6. If the B-subscriber answers, the connection is established.

6.2) CALL FROM PSTN TO MS

The MSC receives its initial data message from the PSTN (via C7) and then establishes the location of the MS by referencing the HLR. It then knows which other MSC to contact to establish the call and that MSC then sets up the call via the BSS serving the MS's location.

7) CONCLUSION

GSM provides recommendations, not requirements. The GSM specifications define the functions and interface requirements in detail but do not address the hardware. The reason for this is to limit the designers as little as possible but still to make it possible for the operators to buy equipment from different suppliers.

In this paper, attempt has been made to name and explained the functions of major components of the GSM network. Efforts have been made to show how the components are connected or designed to communicate over an interface specified by the GSM standards. This inevitably provides flexibility and enables a network provider to utilize system components from different manufacturers (e.g. Motorola BSS equipment coupled with Huawei (NSS)

The article finally gave in detail the MS to Land and Land to MS call sequences to give a good appreciation of the messaging that occurs in the GSM systems, and how the PLMN interacts with the PSTN.

FURTHER READINGS

[1] M. Mouly and M.B. Pautet. “The GSM system for mobile communications’’ published in 1992. ISBN 2-9507190-0-7.

[2] R. Steele ‘’ Digital Mobile Radio Communication’’ Pentech London, Published 1992.

[3] W.C.Y. Lee ‘’Spectrum efficiency in cellular’’ IEEE Trans on Vehicular Technology, Vol. 38. No 2. May 1989.

[4] Ericsson, ‘’ GSM System Survey, 1999.

[5] Motorola Ltd, Principles of Cellular Telecommunications, 1993.

[6] Dr. Hopkins, “ICT as Driving Force of New Economy”, AfriNet 2002.

[7] Dr. F.A. Ibikunle, “ GSM Basics” NITEL Training Journal, Vol. 7, No. 2, Oct 2001.

[8] Motorola, K, and K. Hiradi, ’’GMSK Modulation for Digital Mobile Telephony’’ IEEE Trans. Vol., com-29, No 7, 1981.

[9] Balston, D.M., The Pan- European Cellular Technology. IEEE Conference Publication, 1988.

[10] V.K.Garg and J.E. Wilkes, ‘’Principles and Applications of GSM’’ Prentice Hall, New Jessey, 1999.

[11] Wakid, S., et al., ‘’Coming to OSI: Network Resource Management and Global Reachability’’. Data communications, Dec. 1987.

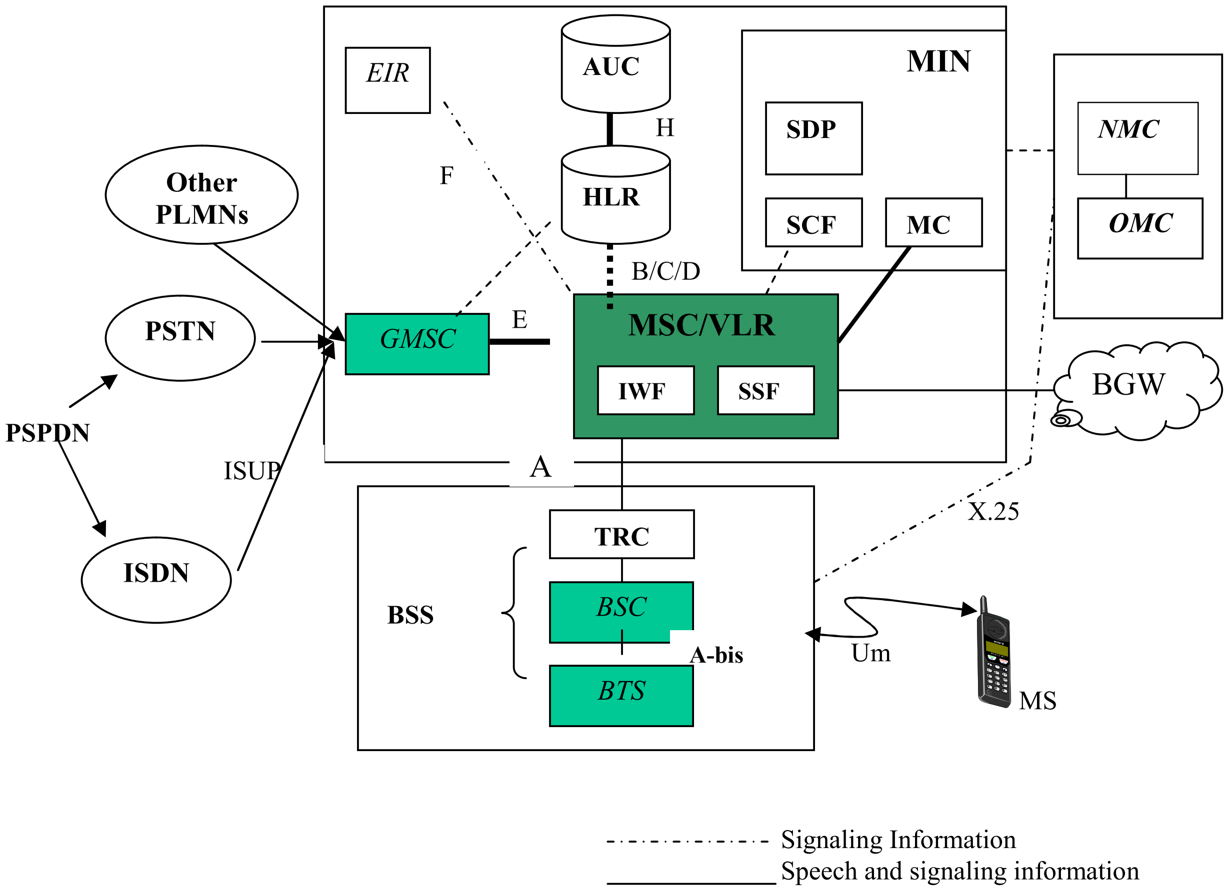


Figure 1: GSM System Overview

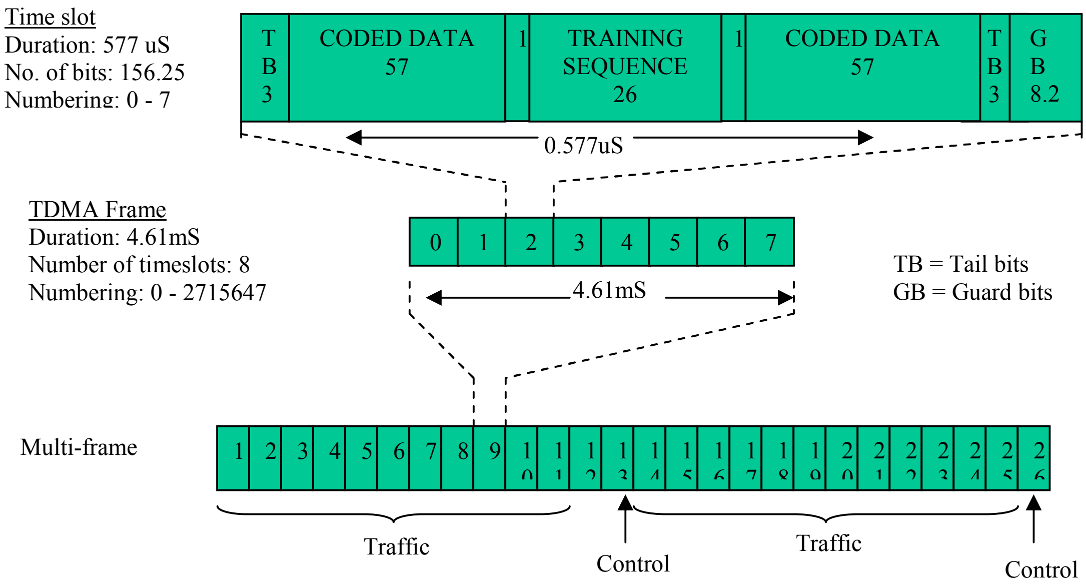


Fig 2: GSM Multiframe, TDMA Frames and Timeslots (Burst)

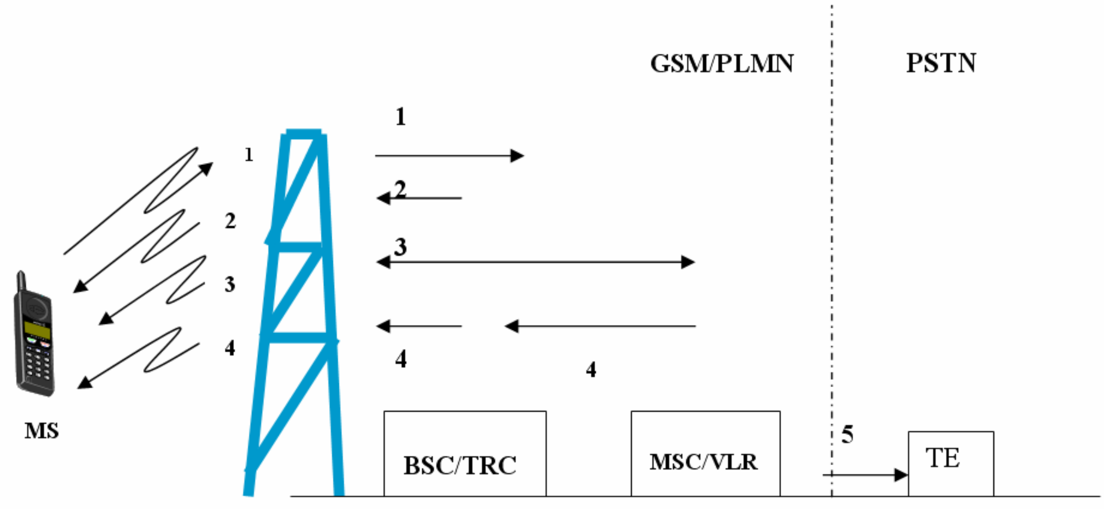


Fig 3: Call sequence from an MS

UNIVERSITY AND THE REVOLUTION IN COMMUNICATIONS AND COMPUTING: EXPERIENCE OF DONETSK NATIONAL TECHNICAL UNIVERSITY

PROFESSOR A. ANOPRIJENKO, E. BASHKOV, AND A. MINAEV

Donetsk National Technical University, UA-83000, Artema, 58, Ukraine

anoprien@cs.dgtu.donetsk.ua

http://donntu.edu.ua

About Donetsk National Technical University (DonNTU)

In the May of 2006 DonNTU will see his 85th anniversary. Our school is the first higher education establishment in the Donbass region and one of the first schools of this rank in Ukraine.

The history of DonSTU could be divided into several stages:

1921-1926 Donetsk Mining College

1926-1935 Donetsk Mining Institute

1935-1960 Donetsk Industrial Institute

1960-1993 Donetsk Polytechnic Institute

since 1993 Donetsk State Technical University

The first in the Donbass Region higher education establishment's coming into being took place under the extremely unfavorable conditions. The mines, metallurgical, machine-building and coke plants were inactive; railways were destroyed. The new government considered the Donbass Region restoration to be a very important task. For this, the engineers were of great demand. There were only 189 mining engineers in the Donbass at that time.

By the beginning of the 90-s DPI had turned up to be the largest higher education school in the Ukraine. More than 1400 academics worked here. Among them there were 122 professors, Doctors of Science, about 700 associate professors, Candidates of Science. 10 Honoured High School Teachers of Ukraine (M. P. Zborshchick, G. V. Maleyev, V. N. Matsenko) and leading researchers who were Honoured by Ukrainian Scientists (V. G. Geyer, V. M. Klimenko, K. F. Sapitsky, N. G. Logvinov, A. A. Minayev and others) trained students. In 1993 DPI was certified and due to the decision of the Ministry of Education, Ukraine obtained the status of state technical university.

Donetsk national technical university is a qualitatively new step in the school's development history.

Every year new specialties and specializations are opened at the University. There are more than 60 of them now. More than 20 thousand students study here. 1300 highly qualified teacher’s trainee students at 88 departments. Many of them are honoured and active members of research associations and academics. They were rewarded with the diplomas of honoured researchers, teachers, and engineers.

For the years of its existence, the University trained more than 110 thousand experts. Political leaders, companies managers, great scientists are among our graduates. These are some of the names: A. F. Zasyadko, N. M. Khudosovtsev, A. A. Pshenichny (they were the Coal Industry Ministers in different periods), A. P. Lyashko -the leader of the government of the Ukraine, E. F. Zvyagilsky -Vice-Prime Minister of the Ukraine, L. F. Bezlepkina-Russia Social Aid Minister. The Donetsk Opera and Ballet House was named after the Honoured Singer A. B. Solovyanenko who was our graduate and teacher. He sang at many world opera stages.

Donetsk national technical university looks into the future. It will occupy a proper place in the Region of mega university training specialists toward the modern technologies of information society in the XXI century.

Development of Internet infrastructure for higher education in Donetsk Region of the Ukraine

The creation of modern high-speed informational and educational infrastructure of Internet in the Donetsk Region is one of the most topical tasks which the regional higher education establishments face. It is connected to the fact that in the nearest future the international community will not recognize the school/people that did not have access to the Internet resources during their training as a specialists with proper education.

The Donbass Region academic infrastructure started its development in 1995 when the first marked channel joined the Institute of Cybernetics of the National Science Academy of the Ukraine and Donetsk Physical and Technical Institute. Donetsk State Technical University and the Regional Library were the first to be on line within the regional Internet network which was being formed by then. This, in fact, initiated the formation of the Donbass Academic and Research Network (DARN) which is the component of the Ukrainian Academic and Research Network UARNet and comprised a number of Donetsk institutes of the National Science Academy of Ukraine and some of higher education establishments of Donetsk. However, the low-grade connection provided rather adequate access only to the E-mail service at that stage.

The formation of the Ukrainian Research and Academic Network URAN being the joint initiative of the Ministry of Education and the National Science Academy was the next stage in the development of the educational Internet in the Ukraine in 1998. The initiative was supported by a number of foreign funds and organizations – the NATO Research Program and the German Research Network in particular. In fact, the beginning of the project network's by Ukrainian Academic and Research Data Network, Part 1 (UARDN-1) implemented by the NTUU Kiev Polytechnic Institute in partnership with a number of leading universities (including DonSTU) and research institutes of the Ukraine was in 1998-1999. The NATO Research Program (grant NIG-97-1779) supported it.

In 2000 the URAN infrastructure was represented by 6 regional centers, DonSTU being one of them (Figure 1). The intensive coordinated development of the regional research and academic infrastructure of the Internet is to be the next stage.

The modern stage of informatization of the educational space of the Donetsk Region requires creation of high-speed universities network aimed to solve the problems they face in a corporate way. There should be three stages in formation of this type of the regional informational and educational network including jointly with Ukrtelecom development of the Eastern node URAN. They are as follows:

Stage 1. Creation of the base digital infrastructure based on backbone with the capacity of 2-100 Mb/s. and peripheral lines (64-128 Kb/s.) covering the leading educational establishments of the region (2001-2002). This stage's activities should be implemented in the following way:

The first stage of the University Informational Ring ("University Ring", Figure 2) which provides for the backbone logging out (100 Mb/s.) to the node of the Donetsk directorate of Ukrtelecom. Besides, it provides for the digital lines to the central node of the network URAN in Kiev for Donetsk National Technical University (DonSTU).

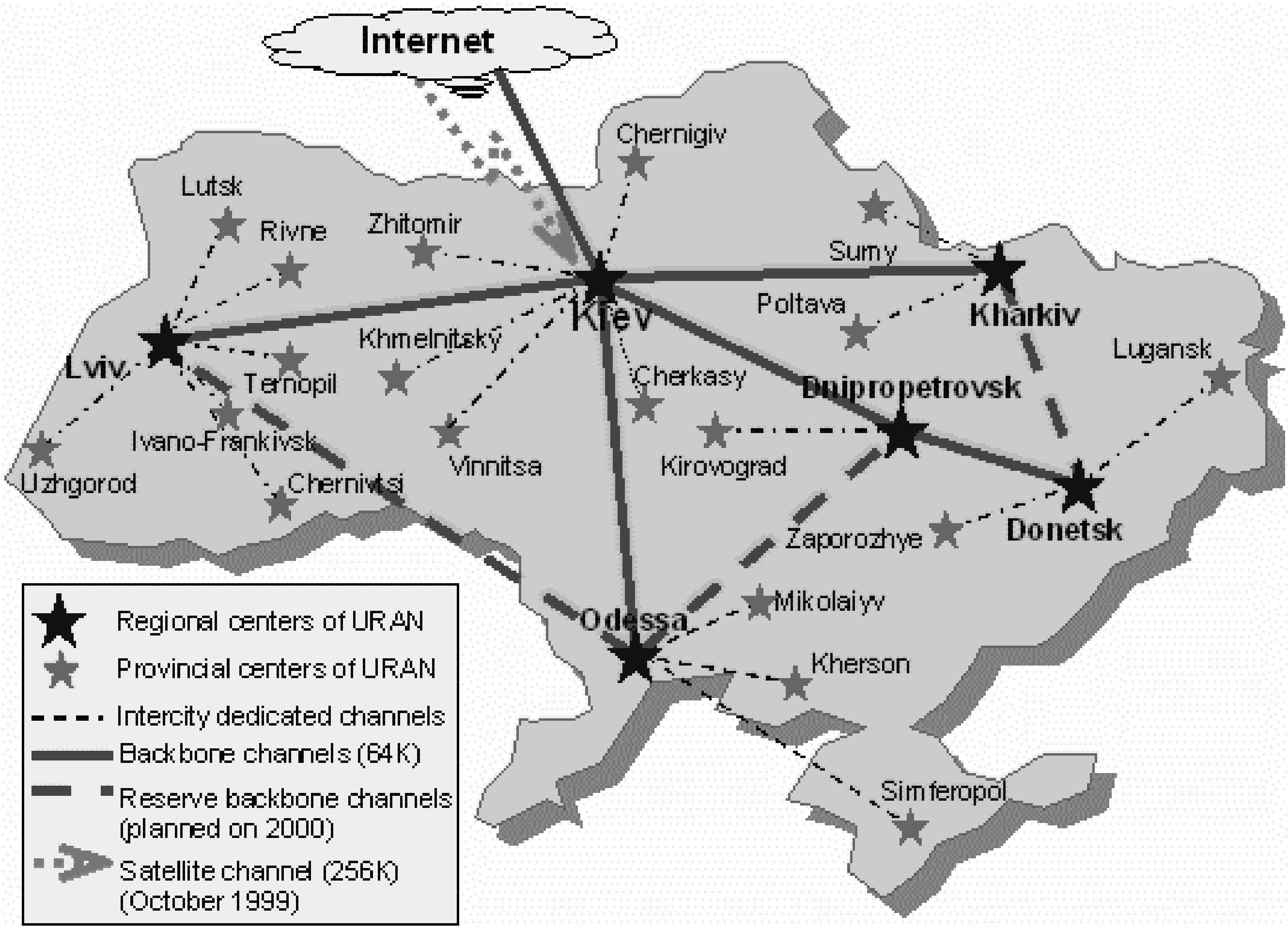


FIGURE 1

Ukrainian Research and Academic Network (URAN): initial 6 regional centers (2000)

Master's portal of DonNTU as the practical result of studying web-technologies

Training students for effective activity in conditions of the European future Knowledge [1] is a complicated task, which requires changes in many approaches to academic process organization. One of the main goals in this case is to make students have the ability to create original, qualitative and attention worthy intellectual product that demonstrates the real level of their knowledge and creativity at their maximum. Master's portal [2] (fig.2), viewed in this report, is one of the results of successful realization of such approach in training the DonNTU students.

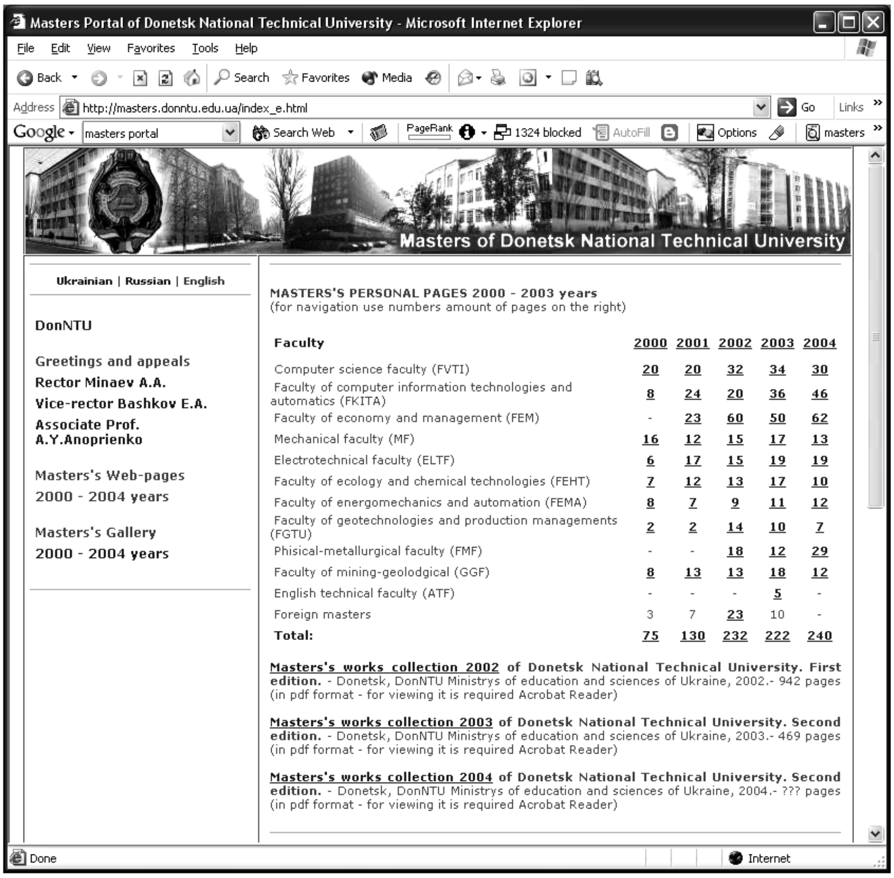


Figure 2 Main english index of a DonNTU master's portal

Systematic work in developing Internet infrastructure possibilities for creation and use of educational-scientific information resources on Donetsk National Technical University (DonNTU) has been conducted since 1996. Characteristic feature of DonnTU approach to presence in virtual reality was initial focus on factuality and multilingual approach (Ukrainian, Russian and English languages with German and French to follow). At present, it is certain that according to information filling and practical use of Internet resources DonNTU continues to be in sure leadership in the region.

Since 2000 DonNTU has been taking goal-oriented steps in expanding and raising quality of its "virtual representation" in the international information world. A considerable intensification of Internet development works in DonNTU was connected with participation in Ukrainian Research & Academic Network (URAN) project [3-5]. The main result of this project (or rather of its beginning phase) in DonNTU became regional server organization of URAN network (http://uran.donetsk.ua), which became the basis for creating a number of educational and scientific resources. DonnTU Masters portal [2] is of the biggest interest, however, as one of the most successful particularized university Internet-projects not only in Ukraine but also in all post-soviet territory.

During the process of search and selection of different projects for information filling of the regional URAN server the crucial part was played by introducing "Network Information Technologies" course into DonNTU Master Program for all specialties in 2000. Quite a number of laboratory works in the given course allowed setting the goal of not only passive study of modern Internet infrastructure information possibilities, but also its active exploration as the extremely effective means of modern scientific communication and electronic publication of educational-scientific oriented materials. The main idea here is that each of Master Program students, while doing this course, will be using knowledge they get not only for fulfilling some abstract tasks, but also for step-by-step creation of their personal site main content of which will be determined by a student's graduation work.

The resource thus formed is raised to such a qualitative level which presupposes possibility of its uploading into Internet and its further practical use. The advantages of such an approach are as follows:

work in mastering network information technologies becomes more purposeful and comprehensive as it presupposes achievement of socially useful information product;

responsibility for results of studying and tasks fulfilment rises as it is supposed that these developed resources will be actively used in Masters portal which is principally different from the usual "archive-trash" use of paper reports;

students become more interested in the results of their work because for the majority of Master students the site, which was developed in the course frames, becomes their first personal Internet page and for quite a while it functions as the so-called "representation in virtual reality";

due to the placement of the web-sites on the Internet as our official resources our students have a chance to master in practice the range of issues connected with correct dealing with copyrights and neighbouring rights, the relevance and observation of which has gained importance in the Internet;

a very important experience in modern terms of cooperative work is acquired in frames of quite scaled and successful Internet-project.

Typical structure of a Master student's site is shown on figure 3. Its basic elements are as follows:

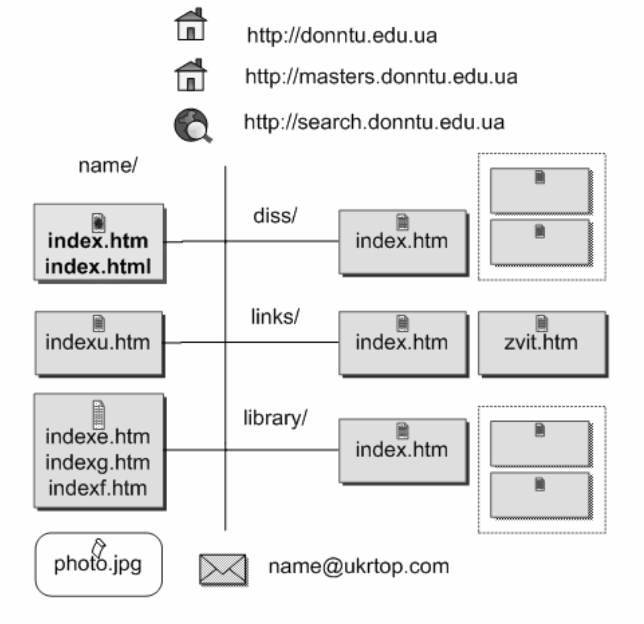


Figure 3 Typical structure of a Master's site

Biographical part includes Master student's Curriculum Vitae (CV) in amount equivalent to 1-3 pages of typescript. Russian (index.htm), Ukrainian (indexu.htm) and English (indexe.htm) versions of CV are obligatory. Alongside with English version, which sometimes is much shorter because of the lack of knowledge of English, German (indexg.htm) or French (indexf.htm) versions are appreciated if a student has studied correspondingly one of these languages. Personal photograph is also obligatory. All the versions of CV should be connected with the mutual hyperlinks and provided with some additional attributes including e-mail of the Master student.

Thematic part consists of three sections connected by the topic of graduation work. The key-section is dissertation abstract (diss/) which is a preliminary description of the main results of the work accompanied by an expanded review on the topic based primarily on materials found in the Internet in the process of the course study. What is more, the last part seems to be more important in the context of the course and that is why it is supposed to be more voluminous than the first one. Hyperlinks section on the topic (links/) is considered to be a necessary addition to dissertation abstract and contains selected annotated links on different Internet resources connected with the topic of a graduation work. The important element of this section is a report on result of thematic search in the Internet (zvit.htm). The third obligatory section is the electronic library on the topic (library/) which contains the most important additional information materials connected with the topic of graduation work or some other creative activity of a Master student including texts of publication with student's participation and some selected materials out of those he/she found in the Internet during thematic search process. Example of index page of a typical Master student site is shown on the figure 4.

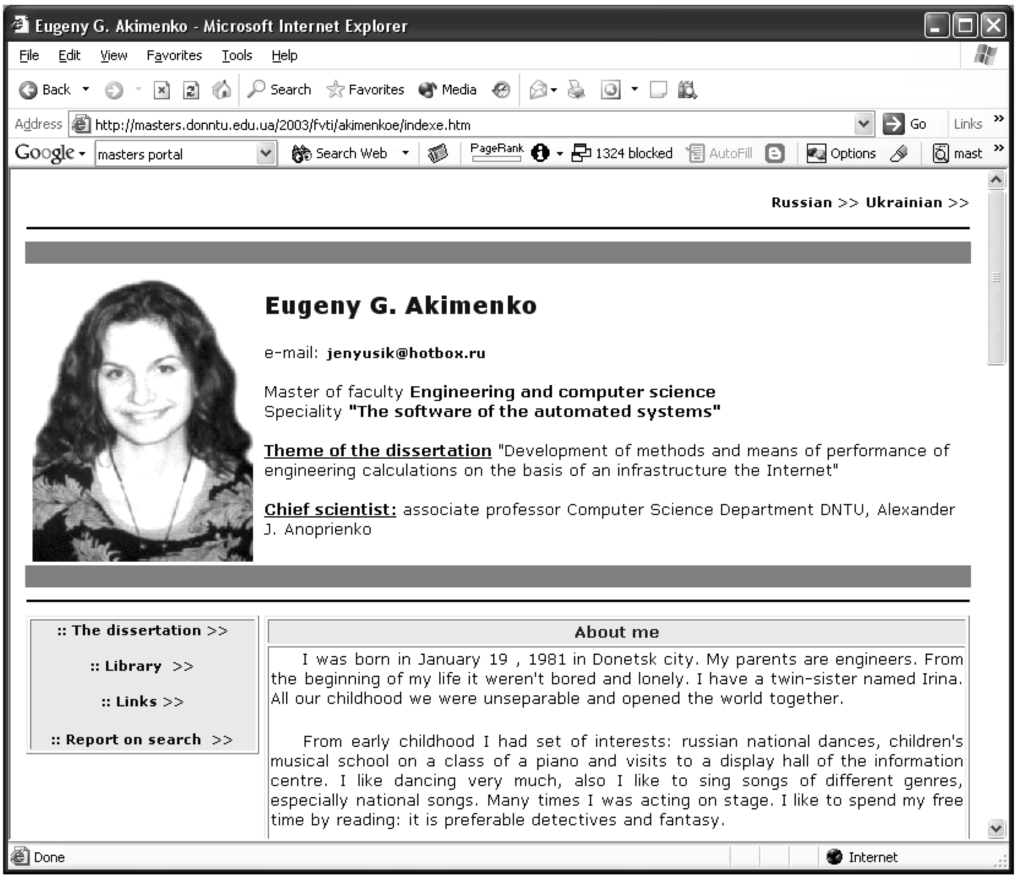


Figure 4 Example of index page of a typical Master's site

Promotion of Masters Portal in the Internet

According to some assessments in the now-forming society not only capital will not be the main resource, but the information itself in its simplest form will not be such either! The factor of "human attention" will be perhaps the determining one, which will be a reflection of the general tendency of shifting such traditional economic and social factors as land, labour and capital to secondary part to the advantage of such specific resource as knowledge. In other words, to effectively train students (and especial Master students) for future it is necessary nowadays, first of all, to train them to use the Internet as an environment of knowledge spreading.

The most objective indicator of value and demand of scientific and educational materials and publications are considered to be the citation index and real intensity of its use. In case of Internet resources it is possible to calculate the activity data with high accuracy and completeness due to special statistic and rating services which was impossible earlier.

Masters portal formation in its present form with the registration in the leading statistic systems has become a reality since the middle of 2002 when there were the results of three Master Program Graduation groups which in overall made up to more than 400 personal thematic sites (Table 1) reflecting practically all the directions DonNTU of research and academic activity.

Table 1

Quantittaive characteristics of information filling

of DonNTU Masters Portal

Year	Number of Master students	Total, Mb	Average site size, Mb
2000	75	22	0,29
2001	130	74	0,57
2002	232	144	0,62
2003	233	148	0,64
2004	240	155	0,65
2005	240	160	0,67
Total	1150	703	

In English-speaking world, the most relevant indicator of a corresponding resource significance is its position in the results of the leading search system Google. Active DonNTU master's portal promotion in English-speaking world has begun since 2003, however, due to initially presupposed multi-lingual approach in the resource, it became possible to occupy quite quickly the leading positions. The proofs for that can be found in the request results in Google and Yahoo for such key words as «master's portal», «masters portal national» (Fig. 5) or «masters portal of technical university».

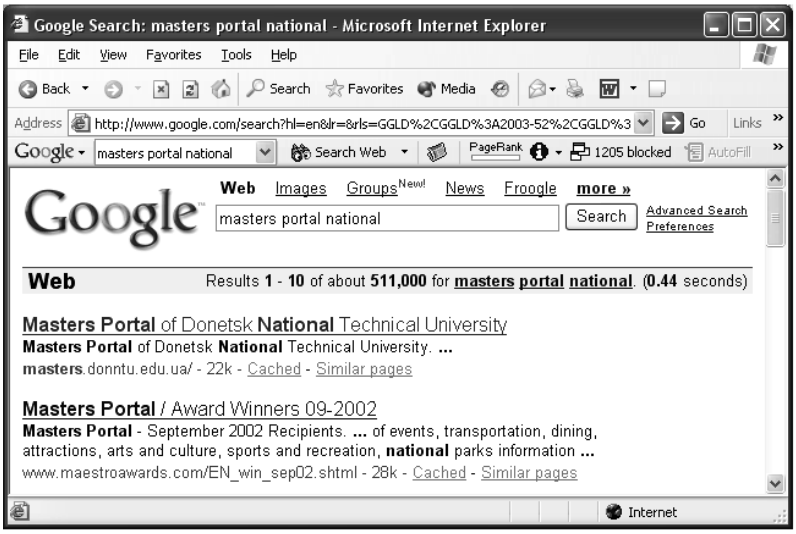


Figure 5 Typical search results in Google for key words «masters portal national».

According to the above said information, special attention was paid to its demand and utility by scientific-educational community. On this way the considerable results have been achieved which can be proved, in particular, by rating positions of the given resource: by 2005 the resource took the leading position among all regional sites, occupied one of the first places in Ukraine in “science” section and entered the first hundred of the most successful ones among all Ukrainian sites.

Figure 6 is quite a demonstration of the fact. It shows attendance changes dynamics of DonNTU Masters portal. The given diagrams are quite typical for research and academic resources: this means that on the background of constant growth of attendance there are pronounced weekly and seasonal rhythms of attendance recession. The weekly maximums are observed in the beginning and/or the middle of academic week, and the seasonal ones – at the end of academic semesters. The attendance minimums happen in weekends and during vacations, in this case indexes decrease almost twice their amount. The absolute minimum takes place approximately in the first half of August. Relatively fast growth in the beginning of the academic year can be explained not only by seasonal demand of the corresponding resources, but also by additional appearance of new personal sites of the current graduation of Master students. This greatly influences the growth of portal summary attendance as its indexation by search engines intensifies.

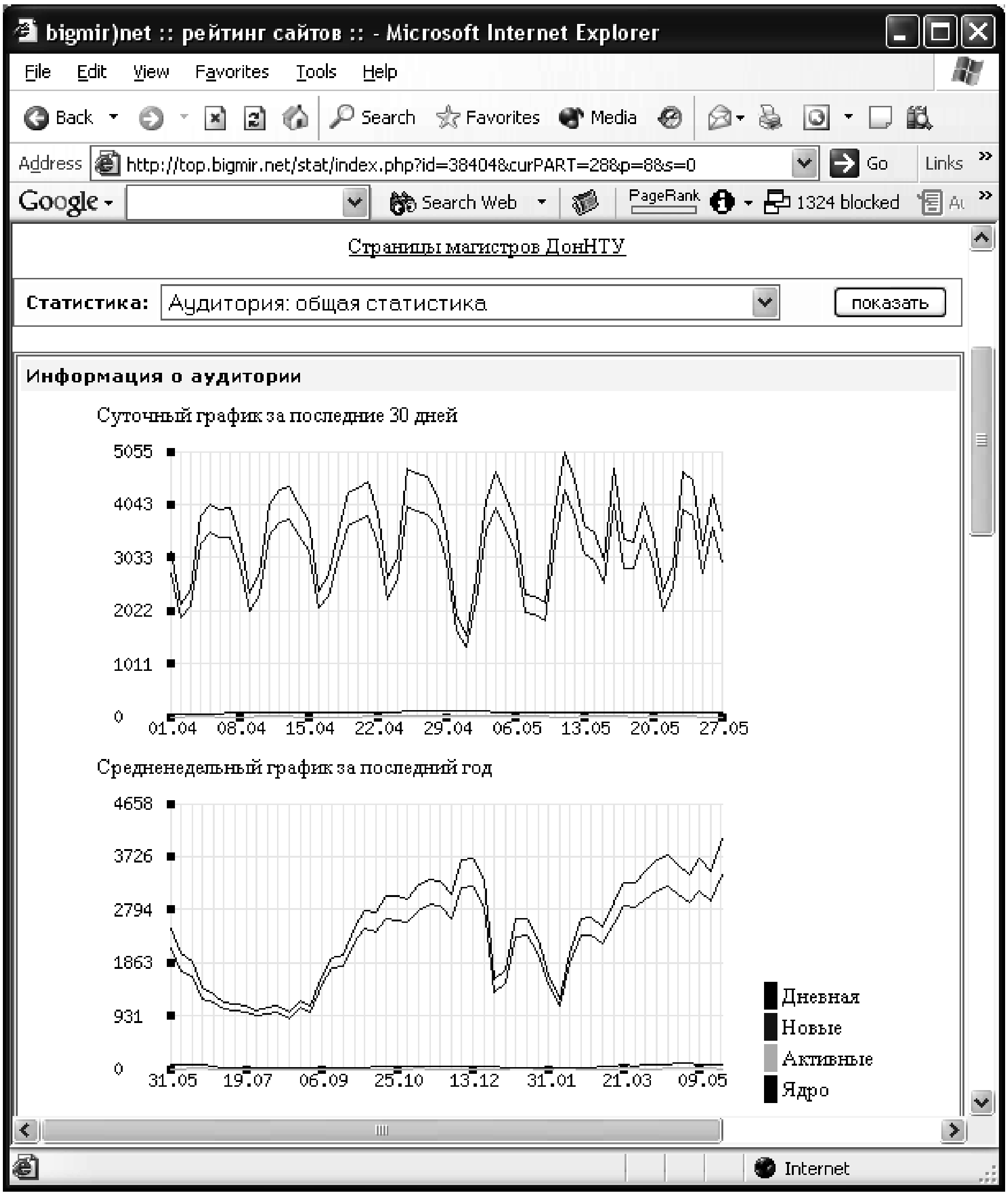


Figure 6 Typicals graphics of visits

Perspectives of further development

Taking into account the five years exploitation experience, the main directions of the portal's further development can be as follows:

Expanding number of participating higher educational establishments and Master students. DonNTU appeared to be actually the first university having created such a resource in all the post-soviet territory. Experience of its successful creation and exploitation lets us hope for appearance of similar resources in other higher educational establishments. One of its main conditions is the respective modern network technologies course included in Master Program's curriculum. In this case optimum alternative would be to borrow DonNTU's experience including compulsory personal standardized web-pages placing all the Master students of the institution or of it's define specialties. It is also possible that some Master students from other different institutions voluntarily place their web-pages in the Internet. Later it is planned to supply DonNTU Masters portal services with such possibility as well.

Expanding the number of information services. Except the above mentioned possibility of independent personal web-pages uploading in the special portal section, it is also appropriate to additionally supply the portal with the Internet resources catalogue on issues connected with Master studies, forum systems, on-line polls and so on.

Portal's methodical element development, in the first place, at the expense of considerable expansion of the number of educational and academic-methodical materials presented there.

Multilingual information filling development of the portal. In particular, German and French versions of portal's generic pages are likely to appear in further years. First of all, however, it is necessary to considerably expand the number of materials in English language which is very important for further resource promotion in the international information world. Development of portal's internal links system, which becomes more and more relevant as information saturation of the resource grows. At present moment in "links" section (links/) of Master students personal sites absolute majority of links leads outside the portal which negatively influences citation index parameters and portal's PR.

Conclusions

On the whole it can be pointed out that DonNTU's experience in creation of the unique academic-science oriented resource and its successful promotion in Slavonic-speaking as well as English-speaking information world is a revealing precedent of a new generation of research-academic resources formation. This portal is also an example of new approaches to academic process organisation using at maximum the information possibilities of modern Internet infrastructure. The main goal for the nearest future is a maximum possible further development of the given resource, and also further search of the ways enabling the most effective mastering of a research-academic potential of modern network information technologies.

References

[1] "The Europe of Knowledge 2020: A vision for University-based Research and Innovation", Liège, Belgium,25-28, April, 2004. <http://europa.eu.int/comm/research/conferences/2004/univ/pdf/conference_proceedings_10-09-04_en.pdf>

[2] Masters Portal of Donetsk National Technical University <http://masters.donntu.edu.ua>

[3] URAN - Ukraine Academic and Research Network (http://www.cei.uran.net.ua/eng/uran.htm).

[4] Present state and development of Ukrainian Research and Academic Networks // NATO Advanced Networking Workshop. Kiev, 2-6 April 2003 <http://www.cei.uran.net.ua/archives/2003apr-workshop/index.htm>

[5] Minaev A., Bashkov E., Anoprienko A., Kargin A., Teslia V, Babasyuk A. Development of Internet Infrastructure for Higher Education in Donetsk Region of the Ukraine // ICEE 2002 Manchester International Conference on Engineering Education, 19-21 August 2002. - Manchester, U.K., 2002.

CONTROL OF THE GSM RADIO RESOURCES

IBIKUNLE FRANK, PHD

M-TEL CHQ., Abuja, Nigeria

1. INTRODUCTION

The frequency spectrum is very congested, with only narrow slots of bandwidth allocated for cellular communications. The number of frequencies and spectrum allocated for various GSM are: Extended GSM 900 (EGSM), GSM 1800 (DCS1800) and PCS1900. A single Absolute Radio Frequency Channel Number (ARFCN) or RF carrier is actually a pair of frequencies, one used in each direction (transmit and receive). This allows information to be passed in both directions. For GSM 900 and EGSM 900 the paired frequencies are separated by 45 MHz, for GSM 1800 the separation is 95 MHz and for PCS 1900 separation is 80 MHz. For each cell in a GSM network at least one ARFCN must be allocated, and more may be allocated to provide greater capacity. The RF carrier in GSM can support up to eight (8) Time Division Multiple Access (TDMA) timeslots, and each RF carrier is capable of supporting up to eight simultaneous telephone calls in theory. In practical, network signaling and messaging may reduce the overall number from eight timeslots per RF carrier to six or seven timeslots per RF carrier, therefore reducing the number of mobiles that can be supported.

Unlike a PSTN network, where every telephone is linked to the land network by a pair of fixed wires, each MS only connects to the network over the radio interface when required. Therefore, it is possible for a single RF carrier to support many more mobile stations than its eight TDMA timeslots would lead us to believe. Using statistics, it has been found that a typical RF carrier can support up to 15, 20 or even 25 MSs. Obviously, not all of these MS subscribers could make a call at the same time, but it is also unlikely that all the MS subscribers would want to make a call at the same time. Therefore, without knowing it, MSs share the same physical resources, but at different times.

On this premises, the paper looked into the factors that determines the efficient use and re-use of the available frequency spectrum in the limited available spectrum of wireless communication systems particularly in the GSM, and concluded with a summary.

2.0 FACTORS GOVERNING THE EFFICIENT USE AND RE-USE OF RADIO SPECTRUM

There are many factors that could contribute and help in the control of the available radio resources in the GSM systems. Enumerated below are the major techniques employed to ensure efficient spectral usage.

2.1) Modulation Efficiency

The modulation technique employed has a direct bearing on efficient use of spectrum. Some highly spectrum efficient techniques developed over the years are 16-QAM, 8-QAM, QPSK and MSK modulation techniques arrange in order of efficiency. But more efficient techniques are usually expensive to implement and may sometime require larger power gains and are used commonly with systems such as high bit rate point –to- point microwave links, as the number of such systems required are small and each of these systems is shared by a number of subscribers. The modulation technique selected for GSM is Gaussian Minimum Shift Keying (GMSK) with a bandwidth data rate product (B.T) of 0.3. The choice of GMSK was based on a compromise of complexity, its provision of a constant transmission amplitude output during the burst and its resilience to co-channel interferences which ensure efficient spectral usage.

2.2) Timing Advance

Timing advance is a solution specifically designed to counteract the problem of time alignment. It works by instructing the misaligned MS to transmit its burst earlier than it normally would. The GSM specifications specify an offset of three timeslots between the Base Station (BS) and Mobile Station (MS) timing, thus avoiding the necessity for the MS to transmit and receive simultaneously. The synchronization of a TDMA system is critical because bursts have to be transmitted and received within the "real time" timeslots allotted to them. The further the MS is from the base station then, obviously, the longer it will take for the bursts to travel the distance between them. The GSM BTS caters for this problem by instructing the MS to advance its timing (that is, transmit earlier) to compensate for the increased propagation delay. This advance is then superimposed upon the three timeslot nominal offset.

In each TDMA frequency band there are eight mobiles, and their transmissions must all be received at the BS in the required TDMA time slot. If, however, one of the mobiles is some distance from the BS then its transmission will be delayed by the time the radio wave must travel relative to the closer mobiles. Despite the fact that the radio waves are traveling at the speed of light, this delay can become significant. Consequently, the BS continuously measures the time offset between its own burst schedule and the reception schedule of the mobile's transmissions The timing advance information is sent to the MS twice every second using the Slow Associated Control Channel (SACCH). In GSM, the timing advance information relates to bit times and an MS may be instructed to advance its transmission by a certain number of bit times. The maximum timing advance is approximately 233µs and at 63 bit times. This is one of the parameters that limit the GSM cell size to a maximum of 35 km radius. Without this timing advance the capacity would be significantly reduced due to the need for increased guard time, as seen for the access burst.

2.3) Sectorization

There are two commonly used methods of sectorization. They are the three 120˚ sectors and six 60˚ sectors, both of which reduce the number of interferences source. Sectorization is used to gain further increase in capacity within a geographical area. It splits a single site into a number of cells, each cell has transmit and receive antennas and behaves as an independent cell. Each cell uses special directional antennas to ensure that the radio propagation from one cell is concentrated in a particular direction. This has a number of advantages: firstly, as we are now concentrating all the energy from the cell in a smaller area 60, 120, 180 degrees instead of 360 degrees, we get a much stronger signal, which is beneficial in locations such as “in-building coverage”. Secondly, we can now use the same frequencies in a much closer re-use pattern, thus allowing more cells in our geographic region which allows us to support more subscribers.

2.4) Cell Radius

Cell radius is perhaps the most important factor governing the spectrum utilization in a wireless system. It plays the role in determining the subscriber density given a certain frequency spectrum. In other words, a smaller cell radius is the key to efficient use of spectrum and one may have to use cell radius as small as 500 meter for a reasonable high subscriber density (Capacity) as will be shown in the work example and the table below. As an example,

Let there be N independent channels available for use in a cell of radius r, and the traffic per subscriber to be e Erlangs. Therefore,

Number of subscribers that can be served in the cell = N/e …………………………(1)

User Capacity (Uc) in a cell ≈ No. of channel in a cell/ Traffic per subscriber x Cell radius

$$= N/(e \times \pi r 2) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Thus, the User capacity is inversely proportional to the square of cell radius. For example, if N = 200, and e = 0.1 Erlang. Then the User Capacity (Uc) will vary with different cell radius as shown in the table using the formula above.

Cell radius(Km)	**User Capacity(Uc)**
r = 25 km	**Uc ≈ 1 per km^2**
r = 10 km	**Uc ≈ 6 per km^2**
r = 3 km	**Uc≈ 70 per km^2**
r = 1 km	**Uc≈ 640 per km^2**
r =500 m	**Uc≈ 2550 per km^2**

Therefore, cell radius plays the dominant role in determining the subscriber capacity given a certain frequency spectrum, and that a smaller cell radius has high subscriber capacity. Small cells are used where there is a requirement to support a large number of MSs, in a small geographic region or where a low transmission power may be required to reduce the effects of interference. Small cells currently cover 200 m and upwards.

2.5) Choice Of Multiple Access

Another key parameter determining the efficient use of spectrum is governed by multiple access technique employed. The access technique defines how the frequency spectrum is divided into channels and affects reuse of the channels.

The oldest technique used in wireless access, especially in mobile communications is FDMA.

2.5.1) Frequency Division Multiple Access (FDMA) have its available frequency spectrum divided in a number of orthogonal frequency channels and these channels are assigned to the user on demand. FDMA can be used both for analog as well as digital communications. This simple technique used extensively in first generation analog mobile system, however, had poor reuse only once in 14 or 21 cells. One way to increase reuse efficiency is by employing sectored or directional antennas at the cell site as discussed in section 2.3. Even with sectorisation, say 3 sectors per cell the best planning gives a typical reuse of once in 7 cells [2], implying reuse actor of 1/7 = 0.143 per cell.

2.5.2) Time Division Multiple Access (TDMA): The most widely access use today for mobile systems and have its available frequency spectrum again divided, but into a few (wide) bandwidth channels or carriers. Each carrier is used for transmission of multiple time multiplexed channels. Each such orthogonal channel (or time slot) could be assigned to a user on demand. This technique's ability to work with smaller signal-to-interference-ratio in digital domain, gives it better reuse factor as compared to the analog FDMA. For example, with 3 sectors, a cell reuse factor of 1/4 or even 1/3 is available [2].

2.5.3) Code Division Multiple Access (CDMA)- A multi-access technique referred to as Direct Sequence Code Division Multiple Access (DS-CDMA) which is based on spread spectrum techniques used extensively in defense applications for over twenty years and it enables definition of near-orthogonal channels in code-space. This access method enables multiple channels to use the same frequency and time slots. Each bit to be transmitted by or for a user is uniquely coded by spreading the bit into 64 or 256 or even 1024 chips. The receiver separate the data of a user by a decoder which correlates the receive signal with the code vector associated with that user. On correlation, the interference from other users would become nearly zero and add only a small amount of noise, where the desired signal will be enhanced considerably. The technique is useful in exploring the inherent time-diversity from multipath delay-spread, especially if the spreading is significant (Chip time of 0.1 μsec to 1 μsec). The only problem with the technique is that as completely orthogonal codes are not possible, especially on the uplink, the total bit-rate supportable from all users using this technique is significantly less than the total bit-rate supportable with TDMA and FDMA technique using the same frequency spectrum. This problem or disadvantage in the CDMA system is made up by better reuse efficiency, as the same spectrum with different set of codes can almost totally be reused in every cell. The theoretical reuse efficiency could be as high as 1.0, but in practice less. With sectored antennas, it is possible to reuse the spectrum in each sector, with a 3-sector cell site resulting in a reuse efficiency of nearly 0.5 per sector.

2.5.4) Multi-Carrier-Time Division Multiple Access (MC-TDMA)– Latest access techniques that has emerged with Dynamic Channel Selection (DCS) [8]. MC-TDMA is a variation of TDMA. A time frame is divided into time slots, as in TDMA; however, in each time slot, a subscriber equipment or Base Station can use any of the several frequencies available. Therefore as in TDMA, the spectrum available is divided into a set of frequencies. Each frequency or carrier can be used in anytime slot by communicating equipment. The key is that no frequency or time-slot is assigned to any subscriber equipment. Nearly all channels are available as a pool for every one to choose from. A technique known as Dynamic Channel Selection governs the choice of the channel and is the key to high reuse efficiency. A reuse efficiency of 0.7 to 0.8 is possible from cell to cell, and with a 3 sector cell, one may get a reuse efficiency of almost 0.7 per sector.

2.5.5)Dynamic Channel Selection (DCS) - Most wireless access systems till recently used Fixed Channel Allocation (FCA) [7] which required a prior allocation or assignment of certain number of channels (carrier frequencies) to a sector in a cell using an exercise generally referred to as frequency planning. The planning had to be carried out using a worst case scenario assuming the nearest possible distance between the interfering signals. Having carried out this worst case planning, having assigned the channel pool to the base station serving the sector, it was up to these base stations (or fixed part) to assign channels to subscribers in the sector on demand.

A totally different approach called ‘’Dynamic Channel Selection’’ does no assignment of channels to any base station or subscriber equipment. All channels are available to every one. The radio equipment is designed to measure the signal strength that it receives on all channels (using something akin to a spectrum analyzer) and thus determine the actual radio environment in its vicinity. It carries out this measurement on a continuous basis, whether it is using a channel or not. Thus the complete knowledge of the radio environment enables it to select a channel in which it can communicate best. The key is that even while it is communicating on one channel, it is measuring the radio environment in all other channels. If it finds that another channel can provide better communication, it switches to this channel seamlessly. Thus the radio equipment (usually associated with the subscriber equipment) continuously monitors all the channels and selects the best channel dynamically. If it succeeds so be it; otherwise, it tries the next best channel.

The DCS is thus based not on worst case scenario, but on actual radio environment. It is thus possible sometimes to reuse a channel even 25m from the other. This reuse even 25m apart is possible because of DCS and gives a reuse factor of 2 to 4 [6] [7]. It is the DCS which gives techniques like MC-TDMA discussed in section 2.5.4 an edge over other multiple access systems that makes it reuse efficiency on the average to be very high.

2.6) Discontinuous Transmission (DTX) / Voice Activity Detector (VAD)

Discontinuous transmission is the process whereby a signal is only transmitted if the user is talking. Whilst they are listening or pausing between words no signal is sent. DTX has two advantages. Firstly, it reduces the power consumption of the mobile by reducing the length of time that transmission takes place. Secondly, it increases spectral efficiency through reducing interference during silence periods. Assuming an average speech activity of 50% and a high number of interferers combined with frequency hopping to randomize the interference load, significant spectral efficiency gains can be achieved. The problem with DTX is that it is often difficult for the MS to detect whether it is speech or noise at its input, particularly when there are high levels of background noise. DTX employs Voice Activity Detectors (VADs) [2] to determine whether speech is present. VAD is a mechanism whereby the source transmitter equipment identifies the presence or absence of speech and its implementation is effected in speech mode by encoding the speech pattern silences at a rate of 500 bps rather than the full 13 kbps These often have a certain delay so that the first 20 ms frame of speech after a period of silence is lost. This problem can be compounded in the case of MS-to-MS calls where DTX is in use on both MS to BS links as this clipping will occur twice, and this so-called double clipping can be highly annoying. For these reasons the use of DTX is limited to one link only.

2.7) Frequency Hopping

In GSM, mobiles are assigned a certain frequency and time slot, and they use this particular channel throughout their conversation. However, this may not always be the case as operators may decide whether they wish to employ frequency hopping. In GSM, frequency hopping takes the form of transmitting each burst on a different frequency to the previous burst. This hopping can improve the spectral efficiency of the system, as frequency hopping tends to reduce the BER, allowing more dense frequency reuse through lower cochannel interference levels. Frequency hopping when combined with interleaving is effective in randomizing the transmission errors which normally occur in bursts. This procedure assists channel coders which work best with a random error distribution. The reason for this is that if the mobile happened to be stationary in a fade during one TDMA burst and if it remained on the same frequency, then it is likely that the mobile would be in a fade during the next burst. If, however, the carrier frequency is hopped to another frequency (beyond the coherence bandwidth) such that the probability of being in a fade is statistically dependent from the fading at the previous frequency, then it is highly probable that the signal will not experience fading. Thus by frequency hopping some error bursts will occur, but they will be relatively rare. On de-interleaving these bursts are randomized.

Frequency hopping leads to another important effect which also acts to increase capacity. This effect is known as interference diversity. If there are two nearby cells assigned the same frequency bands then there is the possibility that they will interfere with each other. Nevertheless, such assignments often need to be made in order to maximize capacity. It may be that in each of these cells, say cell A and cell B, there are four users, users A1 to A4 and B1 to B4 respectively. If say, there were seven frequency bands, F1 to F7 available then user A1 to A4 might be assigned bands F1 to F4 and independently users B1 to B4 might be assigned bands F4 to F7. In this case users A4 and B1, both using band F4 will interfere whereas all the other users will have interference free conversations. In this situation it would be better for all conversations to have a small amount of interference, rather that one conversation having heavy interference. If users A1 to A4 and B1 to B4 are constantly changing their frequencies over the band F1 and F7 in a random and uncorrelated manner, then there will be some bursts when say, A1 will use F4, and thus suffer interference and others when it will use F1 to F3 and not suffer any interference. Since each burst is very short, this effect is averaged out to give the appearance of a lower overall level of interference. Situations where the system would benefit from the interference being more evenly spread happen often and so interferer diversity can have a significant effect.

2.8) Power Control

Control of the transmitted power has two advantages. Firstly, it reduces the power required by the mobile which can continually reduce its power levels to the minimum required. Secondly, it reduces interference by decreasing the overall transmitted power levels. In the GSM system there are eight power categories of base station transmitted power, namely 2.5, 5, 10, 20, 40, 80, 160 and 320W, and five mobile stations given by 0.8, 2, 5, 8 and 20W. Adaptive RF power control is mandatory for the mobile stations, but optional for the base stations. For the MS the output power is controlled in 2 dB steps from a minimum level of 13 dBm up to the maximum transmit power of the mobile class. The MS may change its power level every 60 ms and the time taken for a class 2 mobile to increase its output power from the minimum level of 13 dBm up to the maximum output power of 39 dB would be around 0.78 s. Phase 2 of the GSM specification has introduced four new MS power control levels below the existing ones, allowing ms to reduce its output power to 5 dBm. The output power of a base station transmitter can be reduced from its maximum level in at least steps of 2 dB (with each step accurate to =0.5 dB) to adjust the radio coverage by the network operator.

The adaptive power control is based on RXLEV measurements. In every SACCH multiframe the BS compares the RXLEV readings reported by the MS, or obtained by the base station with a set of thresholds. The exact strategy for RF power control is determined by the network operator with the aim of providing an adequate quality of service for speech and data transmission while containing the interference levels. Clearly, adequate quality must be achieved at the lowest possible transmitted power in order to keep the co-channel interference low. This implies a careful balance in determining the optimum transmitted power. The criteria for deciding on a radio link failure are based on the measurements of RXLEV and RXQUAL, performed by both the mobile and base stations. The procedures for handling radio link failures result in the re-establishment or the release of the call in progress. Down link power control may not be applied to any slots on a BS's beacon frequency. In small capacity cells this seriously reduces the effectiveness of the power control.

2.9) Channel Payload

It is obvious that higher bit rate payload (bandwidth) will require larger frequency resources as compared to a lower bit rates payload. Therefore, for voice communication on wireless systems, it may be desirable to have efficient voice compression and lower bit rate voice codecs. The resulting slightly inferior quality is quite acceptable for mobile communications. But for telephones at homes and offices, toll quality voice communications at 32 kbps/64 kbps may often be desirable.

In GSM, the device that is responsible for these capabilities is the Transcoder. The Transcoder is used to compact the signals from the MS so that they are more efficiently sent over the terrestrial/air interfaces. It reduces the rate at which the traffic (voice/data) is transmitted over the air interface.

2.9) Signaling Overhead

Signaling is a key to setting up calls, monitoring and tearing down of calls. Signaling communications then needs to be carried out on air between subscriber equipment (MS) and the base stations. The signaling channels may be dedicated for each call (user) or may be shared. Usually, the more sophisticated or complicated the system, the more is the signaling requirement. The signaling becomes an overhead that takes away certain frequency resources (bandwidth) and plays a roll in overall efficiency of spectrum usage.

2.10) Call Initialization

Call initialization can occur either due to an incoming call, or due to the mobile wishing to originate a call. If the mobile wishes to originate a call it sends an access message to the BS. If there is an incoming call for the mobile then the BS pages the mobile that then sends an access message. An access request is sent as a message on the RACH. This message is only 8 bits long and does not contain details of the identity of the mobile. The reason for keeping the message short is to reduce the number of access attempts colliding on the random access channel. The BS then broadcasts the channels assignment for that mobile on the Access Grant Channel/Paging Channel (AGCH/PCH).

2.11) Handover

Handover is highly important within a mobile system. Without correct handover the user will have a poor perception of the system as calls may be dropped unnecessarily. Handover also impacts on the spectral efficiency as it may allow overload cells to transfer some calls to neighboring cells with overlapping coverage, and to reallocate frequencies in the case of excessive interference being experienced. However, handovers require BSs to reserve some channels for handover calls and this has an effect on system capacity. The handover process maintains a call in progress as the mobile moves between cells, or when there is an unacceptable degradation of quality caused by interference, in which case an intra-cell handover to another carrier in the same cell is performed. A radio link failure occurs when a call with an unacceptable voice or data quality cannot be improved either by RF power control or by handover. The reasons for the link failure may be loss of radio coverage or very high interference levels. The link control procedures rely on measurement of the received RF signal strength (RXLEV), the received signal quality (RXQUAL), and the absolute distance between base and mobile stations (DISTANCE).

The received signal level measurements are performed on the Broadcast Channel (BCCH) carrier which is continuously transmitted by the BS at a constant level. A mobile measures the received signal level from the serving cell and from the BSs in all adjacent cells by tuning and listening to their BCCH carriers. The root mean squared (rms) level of the received signal is measured for intervals of one SACCH multiframe (480 ms). The received signal level is averaged over at least 32 SACCH frames (≈15 s) and mapped to give RXLEV values between 0 and 63 to cover the range -103 to - 41dBm range in steps of

1 dB. Shorter averaging becomes less reliable but assists faster handovers, which is vital in small urban cells. The RXLEV parameters are then coded into 6-bits words for transmission to the serving via the SACCH. The received signal quality (RXQUAL) is estimated by measuring the BER before channel decoding, using the Viterbi channel equalizer's metrics and/or those of the Viterbi convolution decoder. Eight values of RXQUAL span the BER range of 0.2% to 12.8% before channel decoding.

The absolute distance between base and mobile stations is measured using the "timing advance" parameter explained in section…. The timing advance is coded as a 6 bit number corresponding to a propagation round trip delay from 0 to 63 x 3.69 μs. This allows measurements of BS to MS distances from zero to almost 70/2 = 35 km to an accuracy of about 1 km,

Besides the characteristics RXLEV, RXQUAL and DISTANCE obtained by the measurements highlighted, the radio link control algorithm employs other parameters transmitted by the BS. A mobile need to identify which surrounding BS it is measuring and the BCCH carrier frequency may not be sufficient for this purpose, since in small cluster sizes the same frequency may be used in more than one surrounding cell. To avoid ambiguity, a 6-bit Base Station Identity Code (BSIC) is transmitted on each BCCH carrier in the synchronization burst BS.

4.0) CONCLUSION

We have enumerated in the paper what governs the efficient use and reuse of the available frequency spectrum in the limited available frequency spectrum of a wireless communication system particularly in the GSM. It is evident in the context of this paper, that it is the reduction of the interferences level (in particular, co-channel and adjacent), the power control, modulation efficiency, and the multi-access techniques used are the keys to high spectral efficiency in GSM system. The paper has also shown that without these dynamic procedures, the available resources would not be allocated in an efficient manner and spectral efficiency would deteriorate.

REFERENCES

1. Harfield, D. N., ''Measures of Spectral Efficiency in Land Mobile Radio'', IEEE Trans. Vol 19, Aug. 1997.
2. Lee, W.C.Y., Spectrum Efficiency in Cellullar, ''IEEE Trans., Veh Tech., Vol. 38, May 1989, pp
3. Natvig, J.E., Speech Coding in th Pan- European Digital Mobil Radio System", Speech communication, Special Issue, No. 1, 1988.
4. Murota, K., and K. Hirade, "GMSK Modulation for Digital Mobile Radio Telephony", IEEE Trans., Vol. COM-29, No 7 1981.
5. D.M. Balston, R.C.V. Macario, " Cellular Radio Systems, 1993.

6. Rappaport, T.S., ‘’Wireless Communications – Principles and Practice” Prentice Hall, NewJersey, 1996.
7. Cox, D.C., ‘’Wireless Network Access for Personal Communications’’, IEEE Communications Magazine, Dec 1992.
8. N. Abramson, “Multiple Access in Wireless Digital Access Networks,” *Proc. IEEE*, Vol. 82, No. 9, September 1994.

GLOSSARY

ARFCN Absolute Radio Frequency Channel Number

QAM Quadrature Amplitude Modulation

GMSK- Gaussian Minimum Shift Keying

MSK Minimum Shift Keying

PSK Phase Shift Keying

TDMA Time Division Multiple Access

PART II

TECHNICAL PAPERS: CONFERENCE

COMPRESSING HIGH-DIMENSIONAL DATA SPACES USING NON-DIFFERENTIAL AUGMENTED VECTOR QUANTIZATION

O.O. OLUGBARA[1] AND A.A. ATAYERO[2]

Department of Computer and Information Sciences[1]
Electrical & Information Engineering[2]
Covenant University, Ota, Nigeria
oluolugbara@gmail.com, atayero@ieee.org

Abstract

Most data-intensive applications are confronted with the problems of I/O bottleneck, poor query processing times and space requirements. Database compression has been discovered to alleviate the I/O bottleneck, reduce disk space, improve disk access speed, speed up query, reduce overall retrieval time and increase the effective I/O bandwidth. However, random access to individual tuples in a compressed database is very difficult to achieve with most available compression techniques.

We propose a lossless compression technique called non-differential augmented vector quantization, a close variant of the novel augmented vector quantization. The technique is applicable to a collection of tuples and especially effective for tuples with many low to medium cardinality fields. In addition, the technique supports standard database operations, permits very fast random access and atomic decompression of tuples in large collections. The technique maps a database relation into a static bitmap index cached access structure. Consequently, we were able to achieve substantial savings in space by storing each database tuple as a bit value in the computer memory.

Important distinguishing characteristics of our technique is that individual tuples can be compressed and decompressed, rather than a full page or entire relation at a time, (b) the information needed for tuple compression and decompression can reside in the memory or at worst in a single page. Promising application domains include decision support systems, statistical databases and life databases with low cardinality fields and possibly no text fields.

Keywords: Data Compression, High-Dimensional Data Space, Vector Quantization, Database

1. Introduction

Compression has traditionally not been widely used in commercial database systems because many compression methods are effective only on large chunks of data and are thus incompatible with random access to small parts of the data [1]. Many of the available schemes are only suitable for data compression, which differs from database compression because it is usually performed at the granularity of the entire data objects. In data compression, access to random portions of the compressed data is impossible without decompressing the entire file. Evidently, this is not practical for database systems whose essential function is query processing. Efficient query processing and random accessing to small parts of data without incurring serious overhead is only achievable by fine-grained units like tuple or attributes level decompressions. Compression methods that provide fast decompression and random access are therefore, more attractive for databases than schemes that offer better compression effectiveness. The drawback typically associated with compression is that it puts extra burden on the CPU. However, recent works on database compression [1, 2] have shown that it:

improves system performance especially in read-intensive environments,

provides significant improvement in query processing performance,

reduces disk seek times,

increases disk bandwidth,

reduces network communication costs in distributed applications,

increases buffer hit rate and

decreases disk I/O to log devices.

The problem on which this work premises is in Augmented Vector Quantization (AVQ) [3], which was also called Attribute Enumerative Coding (AEC) [4] or Tuple Differential Coding (TDC) [5, 6]. The goal of the study is to adapt AVQ to address the problem of randomly accessing and individually decompressing tuples, while maintaining compact storage of the data [7]. The original AVQ like many block-oriented schemes such as Adaptive Text Substitution (ATS) [8] compresses and decompresses relation tuples that are locally confined to memory blocks. The problems here are:

block (or page) level compression can result in poor query processing times,

compressed block can cross disk block boundaries and

the size of a compressed block can change when data in a block is updated [2].

Furthermore, random access to individual tuple is still not possible until a block of memory is decompressed. In addition, tuple ordering and differencing in AVQ present overhead cost that are problematic for designing lightweight compression and decompression routines especially when the database is unstable. Both tuple and attributes level compressions were shown to be more attractive from the query processing view point [2]. However, attributes level compression performs better, but has poor compression ratio. The query processing power of tuple level compression, which gives higher compression rate, can be improved upon. This is the motivation for the present work. The solution we propose for randomly accessing stored tuples in a compressed relation is a static bitmap index structure.

The rest of the paper is briefly organized as follows. Section 2 provides the necessary background information on traditional AVQ. Section 3 describes the new scheme and an analogue of bitmap index is suggested for its implementation rather than using expensive B-tree index structure. The evaluation of the method is considered in Section 4 and the paper is concluded in Section 5 with a brief note.

2. AVQ Overview

The AVQ represents a series of tuple values in a relational database $R = < A_1, A_2, \ldots, A_n >$ by the differences between them, where A_i, i=1(1)n are n sets of natural numbers. The method is particularly applicable to sets and databases and it works as follows. First, each tuple in R is treated as integer and R is then sorted by rows. Successive tuples are then differenced and the differences are used to represent R. This technique is formally defined by the quantizer Q_L as follows [3]:

Given a vector quantizer:

$$Q : R \to Z^+,\ Q_L : R \to Z^+ \times N_R$$

is a lossless mapping that encodes a tuple $t \in R$ by the pair <C(t), d(t, Q(t))>, where C is the coder that produces the codeword denoting Q(t) and the difference d between any two tuples is given by Equation (1).

$$d(t_i, t_j) = \begin{cases} \varphi(t_j) - \varphi(t_i),\ if\ t_i < t_j \\ \varphi(t_i) - \varphi(t_j),\ otherwise \end{cases} \tag{1}$$

The compression efficiency of the technique depends on the choice of the codebook. If the codebook is properly designed, the average difference between a tuple and its representative tuple will be small enough that it takes fewer bits to encode than the original tuple. The simplest form of AVQ algorithm is described according to the following steps [3]:

Step 1: Attribute encoding

This is the first preprocessing stage and it achieves compression by mapping a long string of characters attribute to a short number.

Step 2: Attribute domain ranking

The lexicographical order defined by function φ is dependent on the ordering of the attribute domains. Different domain orderings give rise to different orders and different orderings of tuples also give rise to different amount of differences among tuples ordinals, thus affecting the amount of compression.

Step 3: Tuple reordering

Every tuple in R is totally ordered via an ordering rule. The rule is usually a lexicographical order with respect to the attribute sequence in R defined by the function $\varphi : R \to N_R$, where $N_R = \{0, 1, \ldots, \|R\| - 1\}$ and the number $\| R \| = \prod_{i=1}^{n} | A_i |$ is the size of the R space. The function φ is defined for every tuple $t =< a_1, a_2, ..., a_n > \in R$ by:

$$\varphi(t) = \sum_{i=1}^{n} \left(a_i \prod_{j=i+1}^{n} | A_j | \right) \tag{2}$$

The inverse function φ^{-1} defined for all $e \in N_R$ and i = 1(1)n-1 is given by:

$$\varphi^{-1}(e) = < t > \tag{3}$$

Where

$$a_0^r = e, \quad a_n = a_{n-1}^r,$$

$$a_i = \left\lfloor \frac{a_{i-1}^r}{\prod_{j=i+1}^{n} | A_j |} \right\rfloor \tag{4}$$

$$a_i^r = a_{i-1}^r - a_i \prod_{j=i+1}^{n} | A_j | \tag{5}$$

The function φ converts each t ∈ R to a unique integer φ(t) that represents its ordinal position within the R space and a total order is based on this function. To avoid the use of auxiliary variables a_i^r for i = 1(1)n, we replace Equations (4) and (5) by an alternative model (Equation 6) that randomly computes a given a_i from e and previously computed a_{i-j}, j = 1(1)i-1 for all i = 1(1)n. The symbols $\lfloor x \rfloor$ and $\lceil x \rceil$ denote the usual floor and ceiling functions respectively.

$$a_i = \left\lfloor \frac{e - \sum_{j=1}^{i-1} (a_j \prod_{k=j+1}^{n} | A_k |)}{\prod_{k=i+1}^{n} | A_k |} \right\rfloor \tag{6}$$

Step 4: Block partitioning

The ordered relation is partitioned into disjoint blocks of tuples and the size of a memory page is chosen as the partition size. When a tuple is required, the block it resides in is transferred from the disk to the main memory. Coding and decoding of tuples are localized to block level granularity.

Step 5: Block encoding

A particular block consists of a set of ordered tuples and a representative tuple $\bar{t}_k$ is chosen from the block so as to minimize total distortion. Each tuple is therefore, replaced by its difference from $\bar{t}_k$ to obtain numerically smaller tuples with fewer bytes of storage. The leading zero components in each difference are encoded using run-length coding.

3. Non-differential AVQ

The new compression algorithm called Non-differential AVQ (NAVQ) is directly based on AVQ and it compresses a given relation at tuple level granularity. Compression is done while tuples are being stored, thus it supports randomize decompression of individual tuples. The method uses modular arithmetic [9] to associate a tuple t with a pair of integers (q, r), where $0 \leq q < m$ and $0 \leq r < n$. The modular arithmetic partitions the set $\aleph$ of natural numbers into n equivalent classes (or strata), where $n \geq 2 \in \aleph$. Each stratum has $m \geq 1 \in \aleph$ tuples with common features and the entire relation is mapped into a static bitmap index structure $B_{m,n}$. The pair (q, r) serves as the position marker in $B_{m,n}$ for t. We use the term static to discriminate the array from the conventional bitmap index [10], here called dynamic bitmap. The difference between the two bitmap types is that static bitmap is statically created from predefined statistics of the relation and dynamic bitmap is dynamically created from the elements of the relation.

NAVQ is formally defined by a database quantizer Q_N as follows. Given a vector quantizer $Q : R \rightarrow \aleph,\ Q_N : \aleph \rightarrow \aleph \times \aleph$ is a lossless contraction mapping that encodes each $t \in R$ by the pair $< C_1(t), C_2(t) >$<C1(t), where C_1 and C_2 are coders that produce pair of codewords denoting the position of Q(t) in a bitmap. Apparently, $Q_N(Q(t))$ is a mapping composition and the most important distinguishing aspects of the technique are (a) each compressed tuple can randomly be decompressed at a time complexity independent on tuples size, (b) compression and decompression can be carried out without referencing the entire page, let alone the entire relation.

NAVQ is a lossless vector quantization that maps tuple values to bit values. It is different from AVQ since it does not represent a series of tuple values by their differences and does not perform tuple sorting, which can lead to a performance overhead for unstable database. Both methods are similar because they are applied by first treating each tuple in a database table as an integer. Both algorithms use the same mapping φ to convert a tuple to a unique codeword. The compression algorithm is clearly described as follows.

Algorithm 1: DB_Compressor

Input:

The relation R to compress

Output:

A static bitmap structure $B_{m,n}$ of compressed tuples, where $m = \lceil \| R \| / n \rceil$ and $n \in (1, \|R\|)$ are fixed constants.

Method:

This algorithm is basically in phases of preprocessing as follows.

If R is compressible Then

1. Create $B_{m,n}$ and initialize its entries to 0
2. If attribute encoding is required Then perform attribute encoding on tuples in R
3. Map every $t \in R$ to a codeword $e = \varphi(t)$ using Equation 2
4. Map e to $B_{m,n}$ using $B_{q,r} = 1$, where $q = \lfloor e / n \rfloor$ and $r = MOD(e / n)$

End_Algorithm DB_Compressor

The decompression technique is based on the inversion function given by Equation (3) and it works directly in opposite mode to the compression routine. The detail description of the algorithm is given below.

Algorithm 2: DB_Decompressor

Input:

$B_{m,n}$ and n.

Output:

A list of $k \geq 1$ tuples $t_1, t_2, \ldots, t_k$.

Method:

This simple algorithm is the direct inverse of DB_Compressor

If R was compressed Then

1. Create the actual relation R
2. Convert each bitmap entry (q, r) with value 1 to a codeword e using $e = q * n + r$
3. Perform codeword decoding on e using Equation 3 to obtain t $\in$ R
4. If attribute encoding was carried out Then perform attribute decoding on t
5. Insert t into R

End_Algorithm DB_Decompressor

3.1. Application of NAVQ Algorithm

The algorithm was applied to compress the relation R given in [5]. The elements of N_R and the corresponding entries of the bitmap are displayed in Table I. ||R|| = 262144, n = 5 and the efficiency of the technique is 94.44%. Any value of n can be chosen, but for the bitmap to attain high degree of storage utilization, small values are appropriate. We recommend the cardinality of the relation as an appropriate value of n.

Table I: A Relation R

N_R	(q, r)	N_R	(q, r)	N_R	(q, r)	N_R	(q, r)
14816	(2963, 1)	92696	(18539, 1)	154073	(30814, 3)	212130	(42426, 0)
18984	(3796, 4)	100950	(20190, 0)	158233	(31646, 3)	216867	(43373, 2)
21140	(4228, 0)	105118	(21023, 3)	162206	(32441, 1)	223316	(44663, 1)
39331	(7866, 1)	110105	(22021, 0)	173803	(34760, 3)	227484	(45496, 4)
43117	(8623, 2)	117795	(23559, 0)	179038	(35807, 3)	232022	(46404, 2)
47252	(9450, 2)	125352	(25070, 2)	182804	(36560, 4)	235363	(47072, 3)
51104	(10220, 4)	128798	(25759, 3)	186841	(37368, 1)	244658	(48931, 3)
68702	(13740, 2)	134302	(26860, 2)	190996	(38199, 1)	248414	(49682, 4)
80419	(16083, 4)	137827	(27565, 2)	204052	(40810, 2)	252190	(50438, 0)
85140	(17028, 0)	149920	(29984, 0)	207828	(41565, 3)	255449	(51089, 4)

3.2. Data Structure and Operations

We now consider how access mechanisms are constructed on coded tuples and how the tuples can be retrieved and modified. The focus is to give an idea of how the method can be integrated with standard access and retrieval mechanisms. We restrict attention to basic operations rather than to general queries because all queries, simple or complex, reduce to a set of basic tuple operations [5].

We propose a static bitmap index as a suitable data structure for efficient implementation of the algorithms. Dynamic bitmap index is a special kind of index structure consisting in arrays of bits. Each bitmap represents one of the values in the indexed column and the bit position in the array corresponds to the row position in the table [11]. Bitmap indexes are most desirable for low cardinality field relations, systems with low concurrency, few updates and searches. They are frequently used in data warehouses since all of these conditions are found there [11]. A bitmap structure occupies much less space than a correspondent B-tree index [12, 13], an alternative structure that can be used instead of the bitmap. The primary key will then be the q values, which can be made small by choosing n to be large. However, the overhead incur by the q values and pointers to the next node will jeopardize the essence of compression and so bitmap is the most suitable structure for the application since these overheads are avoided.

Static bitmap supports high concurrency, many updates and frequent searches. The structure consisting in arrays of bits, but each array represents a stratum of tuples with common features and not values in the index column. A static bitmap is different from other cached (or pre-computed) access structures in the sense that it does not hold information about relations, but a one–to–one correspondence exists between it and the original relation. The great advantage of static bitmap index is that it allows the database system to avoid direct reading from or writing to the relation. The information needed to answer the query is not taken from the relation and thus in this context, it is related to materialized views [14].

We illustrate the two structures by a concrete example using the Sell relation [15] shown in Table II. The dynamic bitmap indexes and static bitmap index for this relation are respectively given by Tables III and IV.

Table II: Sell Relation

Company	**Product**	**Country**
IBM	PC	France
IBM	PC	Italy
IBM	PC	UK
IBM	Mainframe	France
IBM	Mainframe	Italy
IBM	Mainframe	UK
DEC	PC	France
DEC	PC	Spain

DEC	PC	Ireland
DEC	Mini	France
DEC	Mini	Spain
DEC	Mini	Ireland
ICL	Mainframe	Italy
ICL	Mainframe	France
…	…	…

Table III: Dynamic Bitmap Indexes

(a) Index Company			**(b) Index Product**			**(c) Index Country**				
IBM	DEC	ICL	PC	MA	MI	FR	IT	UK	SP	IR
1	0	0	1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	1	0	0	0
1	0	0	1	0	0	0	0	1	0	0
1	0	0	0	1	0	1	0	0	0	0
1	0	0	0	1	0	0	1	0	0	0
1	0	0	0	1	0	0	0	1	0	0
0	1	0	1	0	0	1	0	0	0	0
0	1	0	1	0	0	0	0	0	1	0
0	1	0	1	0	0	0	0	0	0	1
0	1	0	0	0	1	1	0	0	0	0
0	1	0	0	0	1	0	0	0	1	0
0	1	0	0	0	1	0	0	0	0	1
0	0	1	0	1	0	0	1	0	0	0
0	0	1	0	1	0	1	0	0	0	0

Table IV: Static Bitmap Index

	r_0	r_1	r_2
q_0	1	1	1
q_1	0	0	1
q_2	1	1	0
q_3	0	0	0
q_4	0	0	0
q_5	1	0	0
q_6	1	1	0
q_7	0	0	0
q_8	0	1	0
q_9	0	1	1
q_{10}	0	0	0
q_{11}	0	0	1
q_{12}	1	0	0
q_{13}	0	0	0
q_{14}	0	0	0

3.3. Access method

A static bitmap index structure is constructed using the data in Table (I). Suppose a query wishes to locate the tuple <2, 1, 3, 38, 30>, the query only needs to check if the location given by (32441, 1) for this tuple in the bitmap has a bit value one. The value of one is an indication that the tuple is present in the relation while zero means it does not exist.

3.4. Tuple insertion, deletion and modification

Tuple insertions and deletions are supported in the compressed database as follows. Suppose we wish to insert the tuple <2, 3, 1, 39, 24>. The codeword for the tuple is 186840 and the tuple is stored as 1 in location (37386, 0) in the bitmap structure. For tuple deletion we simply assign bit value 0 to this location. Tuple modification is simply a combination of tuple insertion and deletion. Conclusively, standard database operations remain the same even when the database is NAVQ coded.

4. Evaluation of the Techniques

The performance of a compression technique can be measured in terms of (a) compression ratio, (b) compression and decompression time overhead and (c) query response time. We concentrate on the first factor in this paper since our access structure provides efficient access mechanism.

The efficiency μ of a compression technique operating on a relation R with k tuples is usually defined in terms of two parameters D and C respectively denoting the sizes of the relation before and after compression. The ratio μ is defined by [5]:

$$\mu = 1 - \frac{C}{D} \quad (7)$$

Equation (7) suggests that positive efficiency is not always guaranteed as μ may take on negative values depending on whether the relation is compressible or not. If C > D, negative efficiency occurs and the database is said to be incompressible. A positive efficiency implies the technique compresses well according to the largeness of μ. Small positive value of μ signifies poor compression and zero value shows that compression is not achieved by the technique.

The efficiencies of two compression techniques can also be compared using Equation (7). In this case C and D are respectively the sizes of the compressed relation when techniques 1 and 2 are applied. The value of μ being zero implies that both techniques have the same efficiencies, small value of μ is an indication that the second technique performs better than the first and large value of μ means the first technique compresses better than the second. We further discriminate between two techniques with nearly the same compression efficiency using other characteristics such as compression throughput, simplicity and efficiency of the algorithms. In most cases, a lightweight scheme is preferable to a heavyweight scheme.

4.1 Compression efficiency of AVQ

The efficiency of AVQ is affected by two factors compression overhead per tuple and tuple spacing [5]. The compression overhead per tuple is the size of the count field used to indicate the number of leading zero components of a tuple. Usually, a fixed-size field of size α bits is used to encode this count in order to avoid making the scheme overly complex. For n attribute domains, the number of leading zero components in any difference tuple is larger than zero, but less than n. Thus, the number of bits required for the field is $\alpha = \lceil \log_2 n \rceil$. If relation R has k tuples, the total compression overhead is given by α(k-1), since k tuples yield k-1 differences.

The spacing between two tuples $t_i < t_j$ with respect to φ is measured by a function $\delta : R \times R \to [0,1)$ defined as:

$$\delta(t_i, t_j) = \lceil \log_2(\varphi(t_j) - \varphi(t_i)) \rceil \tag{8}$$

The quantity $\delta(t_i, t_j)$ measures the number of bits required to represent the numerical difference between tuples t_i and t_j. The further apart the tuples are, the larger is the difference. The total space requirement in bits for the k-1 tuple differences and fixed compression overhead per tuple is:

$$C = \sum_{i=1}^{k-1} (\delta(t_{i-1}, t_i) + \alpha)$$

The size of the relation before compression is:

$$D = k\lceil \log_2 \varphi(t) \rceil$$

The compression efficiency ratio of AVQ is given by:

$$\mu = 1 - \frac{\alpha(k-1) + \sum_{i=1}^{k-1} \delta(t_{i-1}, t_i)}{k\lceil \log_2 \varphi(t) \rceil} \tag{9}$$

4.2. Compression efficiency of NAVQ

The efficiency of NAVQ can be determined ahead of compression and it is affected by two factors namely the norm of R space and the overhead of storing the parameter n, which corresponds to column size of the bitmap. Since every tuple is mapped to a bit value, tuple size is just one bit. This follows that ||R|| bits are required for storing a database relation with a maximum of ||R|| tuples. If n is chosen to correspond to the cardinality of R, then the efficiency μ* of the method assuming a low size dictionary was used for attribute encoding, is given by:

$$\mu^* = 1 - \frac{\| R \| + \alpha}{k\lceil \log_2 \varphi(t) \rceil}$$
10)

It can easily be shown that NAVQ gives higher compression ratio than AVQ whenever the load factor of the relation (i.e. the ratio of number of records to the size of the relation) is high. The question of which of the two schemes gives a higher compression rate can be solved as follows. If for a given k, the load factor β satisfies condition (11) then NAVQ gives a higher compression rate than AVQ, otherwise the reverse is the truth. Even if the compression rate of AVQ is higher (for low tuples in relation), NAVQ has the advantage of providing random access to individual tuples in the relation.

$$\beta > \frac{k}{\alpha(k-2) + \sum_{i=1}^{k-1} \delta(t_{i-1}, t_i)} \tag{11}$$

5. Conclusion

This paper presented a new compression algorithm that is based on AVQ and demonstrates its effectiveness on relational database, which exhibits low to medium cardinality fields and numeric fields. The algorithm supports standard database operations, permits very fast random access and atomic decompression of tuples in large collection of data with low decompression cost.

In comparison to a novel AVQ, our technique hopefully yields a higher compression ratio for large tuples. However, in general, the technique has the disadvantage that it compresses only low cardinality field database relations. We hope to develop a hybrid version of this algorithm to compress databases containing generic-purpose data, such as images, sound and text. We also intend to extend the technique for mining association rules in compressed databases.

REFERENCES

[1] Z. Chen, J. Gehrke and F. Korn, "Query Optimization in Compressed Database Systems", ACM SIGMOD, USA, 2001, pp. 271-282.

[2] G. Ray, J.R. Haritsa and S. Seshadri, "Database Compression: A Performance Enhancement Tool", COMAD, 1995.

[3] W.K. Ng and C.V. Ravishankar, "Relational Database Compression Using Augmented Vector Quantization", Proceedings of the 11th IEEE International Conference on Data Engineering, 1995, pp. 540-549.

[4] W.K. Ng and C.V. Ravishankar, "Attribute Enumerative Coding: A Compression Technique for Tuple Data Structures", Proceedings of the 4th Data Compression Conference, 1994, pp. 461.

[5] W.K. Ng and C.V. Ravishankar, "Block-oriented Compression Techniques for Large Statistical Databases", IEEE Trans. Knowledge and Data Eng, 1997, pp. 9, 314-328.

[6] W.B. Wu and C.V. Ravishankar, "The Performance of Difference Coding for Sets and Relational Tables", Journal of the ACM, Vol. 50, 2003, pp. 665-693.

[7] A. Cannane and H.E. Williams, "A Compression Scheme for Large Databases", ACSC. vol. 10, 2000, pp. 241-248.

[8] T.A. Welch, "A Technique for High Performance Data Compression", Computer, vol. 7, No. 6,1984, pp. 8-19.

[9] R.P. Grimaldi, Discrete and Combinatorial Mathematics, An Applied Introduction 2nd ed., Addisson-Wesley, USA, 1989.

[10] C.Y. Chan and Y.E. Ioannidis, "Bitmap Index Design and Evaluation", Proc. SIGMOD Conference, 1998, pp.355-366.

[11] P. Bizarro and H. Madeira, "The Dimension-Join: A New Index for Data Warehouses", SBBD 2001, pp. 259-273.

[12] R. Bayer and E. McCreight, "Organization and Maintenance of Large Ordered Indexes", Acta Informatica1(3), 1979, pp. 173-189.

[13] D. Comer, "The Ubiquitous B-tree", Computing Surveys 2(2), 1979, pp. 122.

[14] N. Roussopoulos, "Materialized Views and Data Warehouses", ACM SIGMOD Record 27(1), 1998, pp. 21-26.

[15] J.G. Hughes, Database Technology, A Software Engineering Approach. Prentice-Hall, 1988.

AN ARCHITECTURAL FRAMEWORK FOR COLLABORATION OF HETEROGENEOUS COMMUNICATION DEVICES USING WAP AND MOBILE DEVICE AUGMENTED (MDA) GATEWAY INTEGRATION

A.A. AZETA, M.O. OYELAMI AND J.O. DARAMOLA

Department of Computer and Information Sciences
Covenant University, Ota, Nigeria.
azeta_ambrose@yahoo.com, olufemioyelami@yahoo.com, wandi_ex@yahoo.com

Abstract

Within the last couple of years, the challenge of displaying collaborative multimedia information has become very important with the large diversity of communication devices such as Personal Computers, laptops, notebooks and handheld devices. The shared data and information may be presented with different views depending on the communication device used by a particular collaborator. The use of various web tools (HTML, WML etc) offers some solutions to the problem but if the target application requires more complex features such as rich multimedia data than is manageable using HTML or WML format, something else need to be done. In this paper, we propose a framework that integrates WAP and MDA Gateway to support collaboration among virtual teams and nomadic workers using heterogeneous communication devices. We then discuss an approach for augmenting mobile device small screen capabilities with surrounding large screen display devices.

Keywords: Collaboration, Heterogeneous devices, WML, Composite Devices, WAP, MDA Gateway

1. Introduction

The growth of the World Wide Web (WWW) in the last couple of years has been enormous and this has made users to embrace the Internet as a standard infrastructure over which a variety of collaboration application and services can be deployed. Collaboration is the process by which teams of people in separate (remote) locations work together, share, discuss and exchange textual and multimedia information. The web is assuming a central role in the way people share information, hence it has been adopted as one of the major media for supporting remote collaboration. For remote collaboration to be widespread, heterogeneous communication devices are often employed, and these are devices such as PCs, Laptops, Notebooks, TV, Handheld devices(cell phones, mobile phones, Personal Digital Assistants (PDAs), pagers, palm top computers, etc) used for exchanging information between remote locations. Increasingly more people and organizations using these devices desire to collaborate in order to improve their communication and productivity. However, one has to take into account that the different users may not be equally equipped in terms of output and input capabilities. Mobile devices often have limited screen display sizes and limited capacity to support audio and video data, and also impose restriction on user input. The restrictions on the input may be due to their small size, miniaturized keyboard (e.g. having keys with multiple functions), sparse capabilities for the capture and recognition of voice and video input. Mobile devices also have limited memory and weak CPU – power when compared to a digital computer. Usability is still a concern with regards to displaying complex graphics on the small screen handheld devices. It is therefore pertinent that something be done to overcome communication barriers imposed by the use of heterogeneous communication devices. Integrating a wide range of communication devices to widen the scope of participants that want to collaborate is a possible solution to these challenges.

Several approaches have been developed to facilitate collaborative systems. In the mobile telecommunication sector, collaboration has been successful with wireless access to the Internet through Wireless Application Protocol (WAP). WAP is a communication protocol and an application environment that enables Internet and web access from wireless handheld devices, with limited display and data capabilities such as pagers, PDAs, mobile phones and other wireless terminals. The WAP gateway serves as an intermediary between the wireless WAP-specific content and WAP client. It also reformats world wide content for display on WAP-based wireless handheld devices. The WAP standard is controlled and developed by the WAP Forum, an industry alliance of more than 200 telecommunication hardware, software, network and peripheral companies[8].

As far as usability is concerned, input, process and display of rich multimedia messages such as photographs, picture messages, maps, electronic postcards, audio and video clips, etc, on small screen handheld device is still a challenge. The capabilities of the handheld device need to be augmented with the surrounding large screen display devices to enhance its functionalities with regards to Multimedia Messaging Services (MMS). The technology of Composite Device Communication Environment (CDCE)[2,6,10] which employs the assistance of the surrounding output devices (Television (TV) set, PCs, etc) to view complex multimedia content can be used. In [6], Composite Device is defined as the composition of available hardware resources that surround user's current location, such as PCs, workstations, high-resolution monitors, TV set etc.

The Multiple Device Augmented (MDA) system being proposed in this paper derives from the CDCE. The gateway that manages the multiple devices (same meaning with composite devices) is the MDA Gateway. The MDA gateway operates as an intermediary between the composite devices, Web server and mobile devices. Depending upon the user's current position and situation, the composite device framework provides a computing infrastructure to incorporate and to outsource/redirect computing tasks to computing or physical resources within the close vicinity. An integrated gateway comprising of WAP and MDA functionalities that support collaboration among users of diverse communication devices is most likely to be quickly and widely adopted as a result of its obvious convenience and flexibility advantage.

This paper proposes a framework that integrate WAP and MDA gateway with surrounding output devices to support collaboration among users of heterogeneous communication devices. It then discusses an approach for augmenting small screen mobile device capabilities with surrounding large screen output devices, to facilitate better display of rich multimedia content.

This remainder of this paper is structured as follows: In section 2, we have a review of related work. The research motivation is discussed in section 3. Section 4 describes the proposed system in terms of architecture and benefits of the integrated WAP and MDA gateway. Finally, a discussion about future research areas and conclusion is given in section 5.

2. Related Work

A number of collaborative systems related to the work described here have been reported in literature. Most of them address collaboration based on textual and non-rich multimedia data using
standard computing devices and mobile devices. Collaboration based on mobile devices and surrounding display devices assisted view, has not been fully exploited.
In [7], a prototype that enables PDA to interact with a TV set was developed. The television responds to PDA output and is used for the presentation of visual images and videos. To overcome a small screen mobile device display constraint, a Composite Device Computing Environment(CDCE) framework was developed in [2,10], it enabled surrounding output devices and computing resources to render and display requested multimedia data. In [6], an ubiquitous computing infrastructure that facilitates nomadic users to access rich multimedia contents using Small Screen/Composite Device (SS/CD) was developed. In [5],
chatting and instant messaging tools were reported. They are tools for keeping people connected and notified. They are targeted towards subscribed users who want to chat in a private network, allowing members of the community to meet, discuss and exchange messages. Some of the players in this field are AOL's Instant Messenger (AIM), Microsoft MSN messaging service and Yahoo Messager. Short Message Service (SMS) also reported in [5] is a very popular service today and has been one of the first messaging tools available for mobile terminal. It is very convenient but has limitations that make it unsuitable technology for rendering collaboration services. These limitations include: 1) unidirectional messaging,
2) a limited fixed length message (160 characters), 3) a message can only be sent from one point to another point. In [3], a framework was proposed for developing adaptive application to the clients computing platform for ubiquitous collaboration. The Collaborator which is a software framework developed by European research project aims at specifying and developing a software distributed environment to support efficient synchronous collaborative work between teams was reported in [4].

The novelty of the approach presented in this work in contrast to existing approaches derives from the integration of an MDA gateway and WAP in facilitating communication of heterogeneous devices especially those within the vicinity of a mobile user.

3. Motivation

We wish to acknowledge the activities of the mobile industry and make the following observations on the future trends of mobile technology.

- Small screen device capabilities will change towards more processing power and high-resolution
  display[6].
- Wireless network and protocol will improve towards more bandwidth. For instance, the third Generations (3G) of wireless network protocols such as General Packet Radio Service(GPRS) and Universal Mobile Telecommunications System(UMTS) offers a much higher bandwidth in comparison to existing previous technologies such as Code

Division Multiple Access(CDMA). The 3G has support for Enhanced Data for GSM Evolution (EDGE) services, a mobile network radio technology that allows GSM network to offer 3G services with existing frequency [9]. The 3G also combine and support
several communication devices-mobile phones, microcomputers and television.

- Due to the requirements of the handheld devices to be pocket-sized, the maximum physical size of the small screen display will remain the same. While it is likely that the handheld device screen resolution and quality will improve, the display size is anticipated to remain constant for a longer period[6].

Being aware of the display limitations and thus the limiting graphical user interface of small screen devices particularly for multimedia data, our focus in this paper is to propose an architectural framework that users of heterogeneous devices can use to communicate using a gateway platform that augment the capabilities of handheld devices with surrounding large screen display devices.

4. Systems Description

In describing the proposed system, we here present the architecture and benefits of the system.

4.1 WAP and MDA Architecture

We have combined the idea in [2] which utilizes arbitrary computing and output devices, such as PC's, TV sets etc, if any is available within the surrounding environment to perform multimedia information/service requests, with tailoring and shrinking capabilities of information of WAP to fit the mobile devices in case no computing or output devices is available within the surrounding environment.

The essential components of the WAP and MDA architecture are(see figure 1):

- WAP Clients: This component is the client device, which includes the wireless handheld devices such as PDAs, mobile phones etc, that a subscriber or a mobile user will use to access the Internet. This client device will have a WAP-based micro-browser(analogous to the desktop web-browser) that will serve as the primary user interface, through which the subscriber will make requests for Internet-based information. Three main and important functions are primarily carried out by the mobile devices. First, the mobile device is used as an interface to access information and services provided by the Web server through the WAP and MDA gateway. Second, the mobile device detects available composite elements(TV, PCs etc) in the close vicinity and informs the gateway. Third, the mobile device task is to control the invoked services and to provide users possibilities to interact with requested information and services.

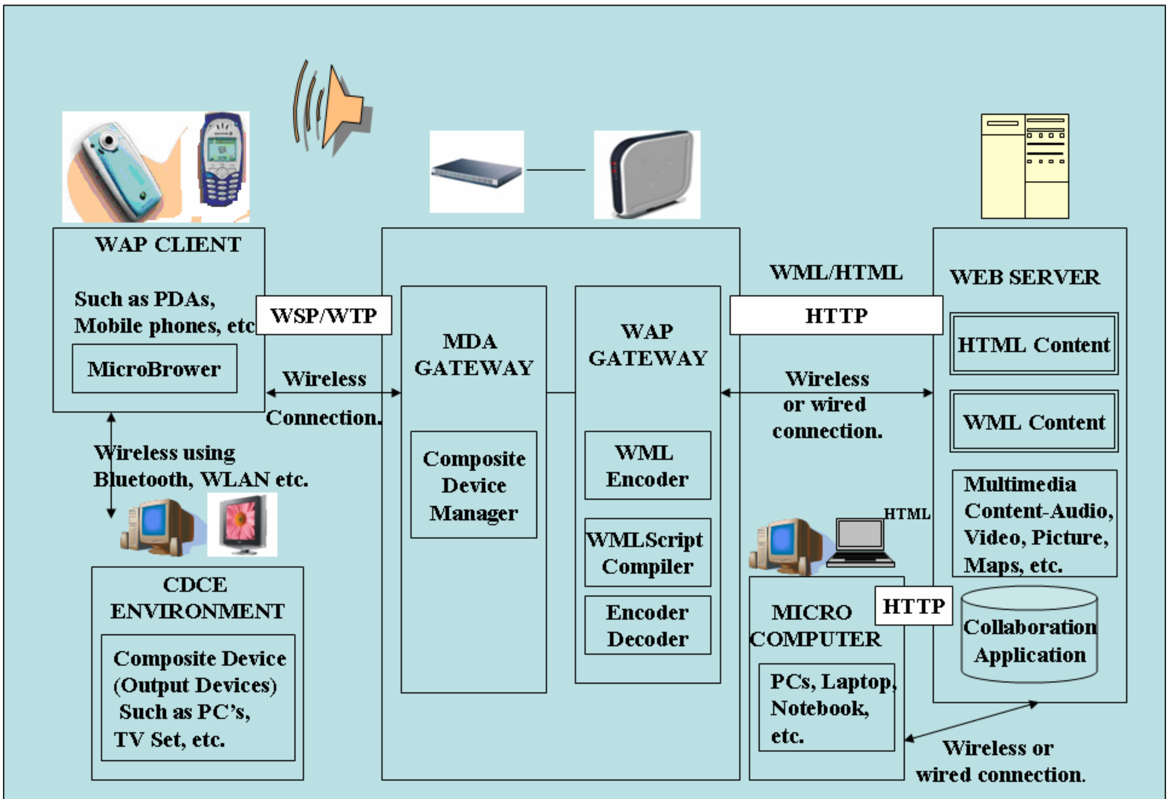


FIGURE 1. A COLLABORATION ARCHITECTURE SHOWING AN INTEGRATED WAP AND MDA GATEWAY

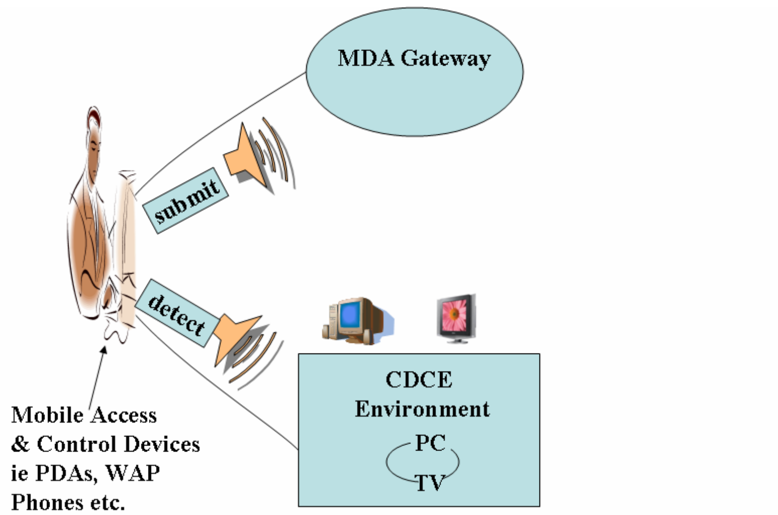


FIGURE 2. AN MDA GATEWAY ARCHITECTURE SHOWING MOBILE ACCESS AND CONTROL DEVICES

- WAP and MDA Gateway: This is one of the important components of the architecture. The gateway is the interface between WAP clients (mobile devices), composite devices and Web server. The functions of the WAP gateway includes: 1) processing of encoded requests for Internet-based information from the mobile devices, 2) decoding/encoding of task, 3) conversion of Wireless Session Protocol (WSP) to HyperText Transfer Protocol (HTTP) and 4) transformation of HyperText

Markup Language (HTML) to Wireless Markup Language (WML) content. WAP also defines a scripting language called WML Script(analogous to Javascript on desktop computers), that extends micro-browser functionality with small applications called scripts. The WML Script compiler does the transformation/translation process from HTML to WML and vice visa.

The composite device manager, the subcomponent of the MDA gateway, manages and maintains the database of available composite elements that can be used to perform certain tasks. Important is the knowledge of the capabilities of each composite element e.g. if a PC has a soundcard and speakers or a monitor has a high color resolution. It assigns the selected services and information request to the most appropriate composite device to perform the task.

- WEB Server: The Web server stores the actual value-added information and content that subscribers wish to access. This content, if in HTML, must be encoded with WML by the WAP gateway for viewing on the WAP clients. The content may also be stored on web-servers in WML, allowing the WAP gateway to pass WML documents directly to the WAP client. Collaboration application may be developed using HTML or WML scripts which may contain textual or multimedia data.

  The operation of the WAP and MDA gateway is such that when a mobile device requests content from the Web server through the MDA gateway and WAP gateway, the WAP gateway retrieves the content directly from the Web server if in WML format, but does reformatting and translation of the content if in HTML format. The requested content will subsequently be returned by the WAP gateway to MDA gateway to determine the nature of the request. For simple output request, the MDA gateway will deliver the content to the small screen mobile device, otherwise it will be redirected to the nearest large screen display device within the vicinity of the mobile user (particularly for rich multimedia request), such as audio/video clip, pictures, maps, photographs, etc.
- Micro Computers: The microcomputer is also one of the clients to the Web server, it uses PCs, Laptops, Notebooks, etc, to access the HTML content of the Web server through HTTP.
- The CDCE Environment: The CDCE comprises of the collection of the output devices

(composite devices) within the vicinity of the mobile user, it includes PCs, TV set etc.

One essential element of the CDCE as seen in figure 2 is a mobile access and control device, which detects the existence of surrounding devices, invokes services on these devices and controls the executed processes. Another distinguished component is the MDA gateway module which is responsible for the management of the composite device and the supply of information and services. The access and control device of CDCE is a mobile device that can be a PDA, WAP phone, etc, which has the ability to detect available surrounding devices as well as to identify its capabilities and availability. The knowledge of the device composition at current location is transmitted to the MDA gateway module that offers different multimedia services to its usage. The user has to submit the selection to the gateway. The gateway assigns the services to appropriate output device for the display and calculates a proper order to execute the services. Optionally, the user can manually influence this process by adjusting the assignment.

The short range communication between the mobile devices and composite devices can be realized using wireless proximity network technology i.e. Wireless Local Area

Network(WLAN), Bluetooth etc. Communication between WAP client and the gateway is wireless, that of the gateway and the Web server could be wireless or wired. The Web server and the microcomputer could communicate through wired or wireless connection.

The WSP specifies compression techniques to provide efficient transmission of the request and response. It also allows for negotiation of capabilities between client and server. While the Wireless Transport Protocol (WTP) is responsible for packet segmentation and reassembly, and for acknowledgement/retransmission of packets[1].

To interact with the system, the methodology used in [4,6] can be employed. The interaction with the virtual mix of computing resources maintained by CDCE as Composite Device using a mobile device is a crucial point of the concept.

To interact with the CDCE environment three phases can be used as follows:

i) The detection phase that determines surrounding output devices.

ii) The construction phase, when the computing resources are constructed to create the composite devices.

iii) The interaction phase, where the user interacts with the invoked services via the mobile device.

Another influence related to the user interaction issues in the interaction phase is the different control modes that can distinguished[6]:

i) Abdicative -The mobile device hands over the control to the output device.

ii) Cooperative - Mobile device and input capabilities of the output can jointly be used to control the application.

iii) Exclusive - The only input device is the mobile device.

4.2 Benefits

An Integrated WAP and MDA Gateway platform will open up new benefits to collaborators and users involved in sharing and exchanging information(including multimedia). It offers:

i) Immense assistance to nomadic workers on the move (i.e. construction and telecommunication engineer, etc), workers on separate sites working and troubleshooting simultaneously on an equipment.

ii) Fast access to business data, for instance, it provides doctors and nurses with the ability to access patient data at any time while being mobile within the hospital vicinity. This kind of communication infrastructure facilitates and provides a better point of care services at patient's bed[6].

iii) Enormous assistance to users on mobile commerce and other business transaction purposes ie a prospective home buyer could use a Personal Digital Assistant (PDA) to access a home information service on an Interactive Television(ITV)[7].

iv) Effective facility for video conferencing and virtual cooperation.

v) Deployable ubiquitous computing environment in which different kinds of computers and devices surrounding users are engaged to provide computing services.

vi) Mobile users access rich multimedia content and services without having to shrink or to tailor the content to match the capabilities of the mobile devices.

vii) Users access to rich multimedia contents and services without having to compromise much in quality and diversity.

viii) The creative activities of collaborators are enhanced due to the availability of bigger size output device.

5. Conclusion

In this paper, we have proposed an architectural framework to facilitate virtual teams and nomadic workers to collaborate and exchange textual and multimedia content. The integration of WAP and MDA gateway facilitates an enhanced view of information. The MDA component of the gateway allows complex multimedia data requested by mobile devices to be redirected to the nearest surrounding large display devices for better view. Other issues connected to security and privacy can be addressed in future work.

References

[1] J. L. Berthou, Wireless Application Protocol(WAP), Available: http://www.e.kth.se/~x99_jlb/wap.html

[2] G. Schneider, S. Djennane, T. L. Pham, and S. Goose, Multimodal Multi-Device UIs for Ubiquitous Access to Multimedia Gateways, Proceedings of the 17th International Joint conference on Artificial Intelligence: Workshop on Artificial Intelligence in Mobile Systems (AIMS), 2001.

[3] A. M. Krebs and I. M., Adaptive Applications for Ubiquitous Collaboration in Mobile Environments, Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 1 p. 10031c

[4] F. Bergenti, S. Costicoglou and A. Poggi. A Portal for Ubiquitous Collaboration. The 15th Conference on Advanced Information Systems Engineering(CAiSE Workshop 2003), Klagenfurt/Velden, Austria, 16-20 June, 2003, Workshops, Proceedings, Information Systems for a Connected Society.

[5] S. Djennane, S. Goose, WAP collaboration for mobile users http://www9.org/w2-mobileweb/WapChat.htm.

[6] T. L. Pham , G. Schneider, S. Goose, A situated computing framework for mobile and ubiquitous multimedia access using small screen and composite devices, Proceedings of the eighth ACM international conference on Multimedia, p.323-331, October 2000, Marina del Rey, California, United States

[7] S. Robertson, C. Wharton, C. Ashworth, M. Franzke, Dual device user interface design: PDAs and interactive television, Proceedings of the SIGCHI conference on Human factors in computing systems: common ground, p.79-86, April 13-18, 1996, Vancouver, British Columbia, Canada.

[8] S. N. Parekh, A Closer look at the Wireless Application Protocol (WAP) http://itc.mit.edu/itel/students/sohil/ParekhWAP_0500.pdf

[9] C. K. Ayo, Phd lecture notes on Computer Communications Networks and distributed Processing, 2006, (unpublished).

[10] T. L. Pham, G. Schneider, S. Goose, A. Pizano, Composite Device Computing Environment: A framework for augmenting the PDA using surrounding resources, Workshop on Situated Interaction in Ubiquitous Computing at CHI2000, Germany, April 2000.

A PROCESS FRAMEWORK FOR SUBSCRIBER MANAGEMENT AND RETENTION IN NIGERIAN TELECOMMUNICATION INDUSTRY

J.O.DARAMOLA, O.O. OLADIPUPO AND A.G. MUSA

Department of Computer and Information Sciences,Covenant University, Ota, Nigeria.

{wandi_ex, frajooje, bola_musa }@yahoo.com

Abstract

Churning which is a sudden defection of subscriber to competitors is a disturbing problem in the global telecommunication industry. Hence, a dominant approach for subscriber management and retention is churn control, since it is cheaper to retain an existing subscriber than acquiring a new one. Predictive modeling employs the use of data mining techniques to identify patterns and provide a result that a group of subscribers are likely to churn in the near future. However, the effectiveness of subscriber retention strategy in an organization can be further boosted if the reason for churn and the timing of churn can also be predicted.

In this paper, we propose a data mining process framework that can be used to predict churn, determine when a subscriber is likely to churn, provides the reason why a subscriber may churn, and recommend appropriate intervention strategy for customer retention using a combination of statistical and machine learning techniques. This experiment is carried out using data from a major telecom operator in Nigeria.

Key words: Churn prediction, telecommunication, decision support system, survival analysis, artificial neural networks.

1. Introduction

The competitive nature of the wireless telecommunication industry derives from having several wireless telecom operators striving to win over a particular population with services that are basically similar. This often leads to release of new services, promotion of incentives packages, endless advertisements and incessant new rate regimes. The effect of this, especially for a fully matured market is the tendency for a subscriber to migrate to another competitor. Although, the Nigerian wireless telecom market is not fully matured with less than 20% of the entire population as wireless subscribers, the elite population market is fully matured, and so churn control among this high-valued subscriber population should be of uppermost concern to any proactive telecom operator.

The phenomenon of churning is a process of defection of a subscriber from a company to its competitor. Churning is a major market repositioning situation which usually occurs when the market in which two or more companies compete, matures. When it comes, the whole organization has to change. It not only places a concern on a product, a person or a department but it is a challenge at all levels of the entire company hierarchy. To handle this, the company has to take a holistic view on the cause of the churning and what to do about it.

A general customer base analysis places an emphasis on effective churn management. Churning should be well understood by taking a cursory look at the phenomenon from the competition, the market and the external factor point of view. Of importance is the statistical ratio of customers loss to recently acquired customers. A high churn rate is a pointer to the fact that customers are dissatisfied with the quality of services rendered and a low or near competitor rate indicates the market is highly competitive and growing. When the rate is high, drastic measures should be taken to bring forth reasons for churning and strategies for subscriber retention.

Quite a number of models and techniques have been used to predict customer churn in literature. A large number have extremely delve into churn prediction and quite a number have extended the prediction to include when a customer is likely to churn and the reason(s) why a customer might churn [10],[11],[18]. Also, some of the techniques that have been used to predict customer churn includes: logistic regression [13], decision tree (Classification And Regression Tree - CART) [2], artificial neural network [3], [17], data mining [7],[14], boosting [13], and genetic algorithm [1].

In [13] Mozer C.M. et al. presented a scheme for predicting subscriber dissatisfaction and improving customer retention and profit maximization using statistical machine learning techniques, their work did not address the issues of why and when churn will likely take place. [5], is a case study report of churn prediction in telecommunication data using of a data mining software package called MiningMart. The focal point and critical success factor explored in this approach was a clever preprocessing of the given data. In [6] the capabilities of three predictive models at predicting churn namely: neural network, logistic regression and decision trees were compared using complaint data. The aim was to identify the most suitable model hence a statistical analysis was done on each of the models in terms of their predictive accuracy. In [8] was the report of a survey of churn and customer loyalty in the Korean telecom market.

Our approach in this work is not only to build a model that can predict which subscriber may churn, but with extended capabilities to advance reason(s) why subscriber may churn, when they may churn and recommend appropriate strategy for their retention. Therefore, our process framework handles churn prediction from four perspectives. We intend to implement this framework from a rich chunk of telecom subscribers' demographic data, subscribers' transactions information and subscribers' complaints information.

The outline of the other aspects of this paper is as follows: section 2 is a description of the process framework, in section 3 we gave a description of the experimental procedure with prototype data. In section 4 we discuss our result and in section 5 we have the conclusion.

2. The Process Framework

We discovered that many of the approaches reported in literature occur in contexts where the detailed information about subscriber like demography and transaction information are available, example is corporate clients. A dominant group of subscribers is the prepaid subscribers, where limited demographic information have been made available (an example is the Nigerian context), yet this group of subscribers forms the greater bulk of an operator's subscriber base. Our effort in this work is focused on building an appropriate prediction model from available information on prepaid subscribers. The process framework being proposed will facilitate data mining of telecommunication data such that it is possible to predict churn, adduce reasons for a possible churn and predict when churn will take place.

The figure below (see Figure 1) shows the framework architecture for subscriber management to facilitate churn prediction, predict when a customer is likely to churn, provide the likely reasons for a churn and give recommendations of intervention strategies for subscriber retention. Data from a telecom service provider about its subscribers comprising of the complaint and repair data, and customer transaction data will be used to build an appropriate data prediction model. The data model is then processed for churn prediction, and survival analysis. Thereafter, the results of the two analysis procedures are passed to a decision support expert system component that recommends the most appropriate intervention strategies for customer retention.

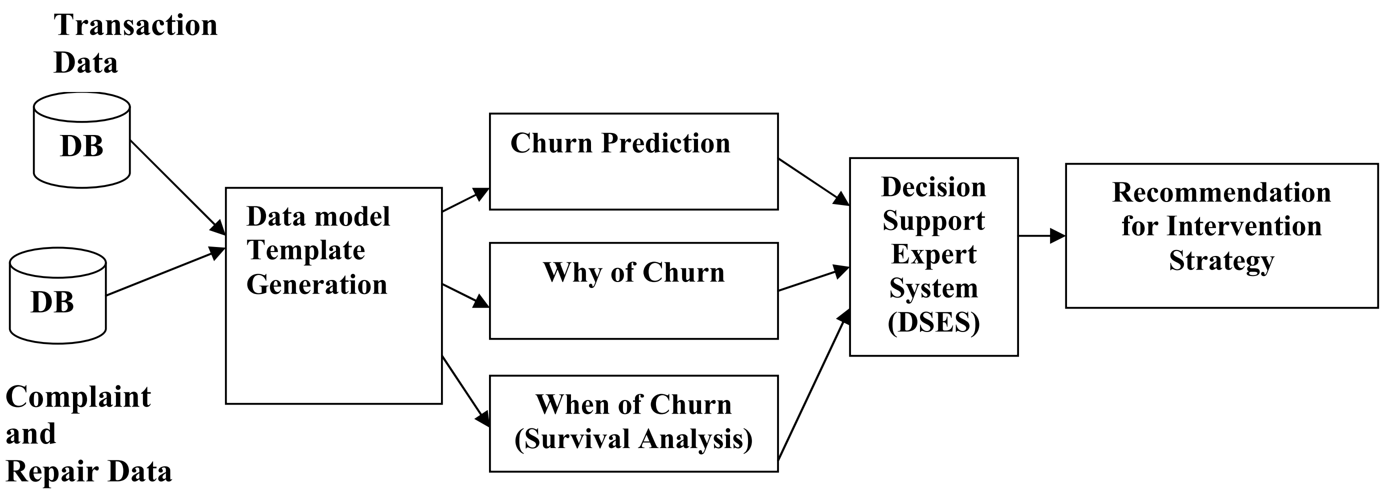


Figure 1: Framework Architecture for Subscriber Management and Retention

2.1 The input Data Set

The subscriber data used for our experiment was gleaned from the schema of a major wireless telecom operator (as the operator will not want to be identified, as churn prediction results are confidential). Although relatively few real data had been given, we were allowed access to the schema of the relevant databases. The schema of the subscriber transaction database is given as follows:

Subscriber = (*caller_number*, *called_number*, *incoming_route*, *outgoing_route*, *amount_b4_call*, *amount_after_call*, *International Mobile Subscriber Identity* (*IMSI)* (unique), *exchange_id*, *record_type* (sms, voice, gprs etc.), *event_type*). A typical subscribe transaction can be selected from the transaction database using appropriate query statement such as:

*"**Select** call date, details, caller_number, called_number, call_duration, call_charge, balance_before, balance_after **from** transactionDB **where** caller_number = subscriber_Number".*

The complaint data containing the records of complaint by subscribers has the structure:

Complaint_data = (*request_complaint_id* (unique), *date of complaint*, *time of complaint*, *type of complaint*, *status* (open or closed), *imputer* (internal staff initiator), *handled_by* (person to whom the problem was assigned)). The nature of the complaint database and subscriber transaction allows for multiple occurrences of a records of transaction and complaints of subscribers. Other relevant information about a subscriber include: *date of subscription, type service subscribed, subscriber number (unique)*

2.2 Data Prediction Model

The parameter template for predicting churn and the reason for churn were extracted from subscriber transactions and the complaint records datasets. A typical subscriber complaint falls into one of the established categories that have been denoted by unique codes, these are: swim swap billing (ssb), damaged recharged cards (drc), inabilibity to load credit (ilc), inability to make or receive calls (imr), roaming problems (rb), migration issues (mi), gprs (gp), wasp (wp), vas (vs) , dealers (dl), handset problems (hp), vip issues (vp) etc.

The parameters template as extracted from the input datasets used in our experiment is presented as follows:

1. Calling number (unique)
2. Calls made ratio = (total number of calls made / number of days)
3. Calls received ratio = (total number of calls received / number of days)
4. Calls dropped ratio = (total number of incomplete calls / number of days)
5. Amount charged ratio = (total amount charged / number of days)
6. Service utility ratio = (total calls duration / total number of calls)
7. Credit load ratio = (total credit load/ number of days)
8. ssb ratio = (number of ssb complaints/ total number of complaints)
9. drc ratio = (number of drc complaints/ total number of complaints)
10. ilc ratio = (number of ilc complaints/ total number of complaints)
11. imr ratio = (number of imr complaints/ total number of complaints)
12. rb ratio = (number of rb complaints/ total number of complaints)
13. gp ratio = (number of gp complaints/ total number of complaints)
14. wp ratio = (number of wp complaints/ total number of complaints)
15. vs ratio = (number of vs complaints/ total number of complaints)
16. dl ratio = (number of dl complaints/ total number of complaints)
17. hp ratio = (number of hp complaints/ total number of complaints)
18. vp ratio = (number of vp complaints/ total number of complaints)
19. Complaint frequency ratio = (number of complaint / number of days)
20. Response ratio = (number of closed complaints/ total number of complaints)
21. Mean-response = (total time spent on closed complaint / number of closed complaint)
22. No-response ratio = (total number of open complaint / total number of complaints)

3. The Experiment

3.1 Churn Prediction with ANN

We implemented an algorithm to scan the input datasets (i.e. transaction and complaint data) in order to generate the values for the parameters of the prediction template. Table 1 shows sample data values after the generation process. The values in the prediction data model were rescaled using interval scaling with the formula:

$$\mathbf{A_i = (V_i - V_{min}) / (V_{max} - V_{min}) \quad For\ i = 1 \ldots n} \qquad (1)$$

The rescaled values were used as inputs into a feed-forward back propagation artificial neural network model (see Table 2). The network outputs are also unscaled using the formula:

$$\mathbf{V_i = A_i (V_{max} - V_{min}) + V_{min} \quad For\ i = 1 \ldots n} \qquad (2)$$

The churn prediction neural network model was implemented using MATLAB 7.0 Software [12]. The configuration of the feedforward neural network model used was a 21-22-1 Multi-Layer Perceptron (MLP) [4], with 21 neuronode units in the input layer which directly corresponds to the number of input parameters in our prediction model template, 22 neuronode units in the hidden layer (1) and one neuronode unit in the output layer that corresponds to the output (1 or 0) which is indicative of whether a subscriber churns or not. The Levenberg-Marquardt (TRAINLM) backpropagation algorithm was used in training the network with the following training parameter values:

net.trainParam.epochs 500	*'Maximum number of epochs to train*
net.trainParam.goal 0	*'Performance goal*
net.trainParam.max_fail 5	*'Maximum validation failures*
net.trainParam.mem_reduc 1	*'Factor to use for memory/speed tradeoff*
net.trainParam.min_grad 1e-010	*' Minimum performance gradient*
net.trainParam.mu 0.001	*' Initial Adaptive learning parameter value (Mu)*
net.trainParam.mu_dec 0.1	*'Mu decrease factor*
net.trainParam.mu_inc 10	*'Mu increase factor*
net.trainParam.mu_max 1e12	*'Maximum Mu*
net.trainParam.show 25	*'Epochs between showing progress*
net.trainParam.time inf	*' interval*

Figure 2 show the graphical representation of pattern of weight convergence and number of training epochs.

080x2	3500/182	2650/182	56/ 182	75400 /182	15600/ 3120	56800/ 182	65/ 700	160/ 700	80/ 700	75/ 700	83/ 700	45/ 700	90/ 700	40/ 700	15/ 700	39/ 700	8/ 700	700/ 182	600/ 700	2006/ 600	100/700	1
080x3	2560/100	160/ 100	30/ 100	89000/ 100	10000/ 2720	6800/ 100	40/ 200	10/ 200	30/ 200	15/ 200	8/200	7/ 200	2/ 200	50/ 200	20/ 200	10/ 200	8/ 200	65/100	63/65	900/63	2/65	0
080x4	5000/365	2010/365	20/ 365	125005/365	18955/ 7010	120000/365	150/ 850	72/ 850	45/ 850	99/ 850	80/ 850	40/ 850	65/ 850	100/ 850	40/ 850	9/850	150/850	950/ 365	950/ 950	1960/ 950	100/950	0
080x5	4560/270	4000/270	10/ 270	56892/ 270	10690/ 8560	50565/ 270	100/ 1200	110/ 1200	70/ 1200	5/ 1200	280/ 1200	63/ 1200	79/ 1200	95/ 1200	106/ 1200	215/ 1200	77/ 1200	550/ 270	450/ 550	4698/ 450	100/550	1

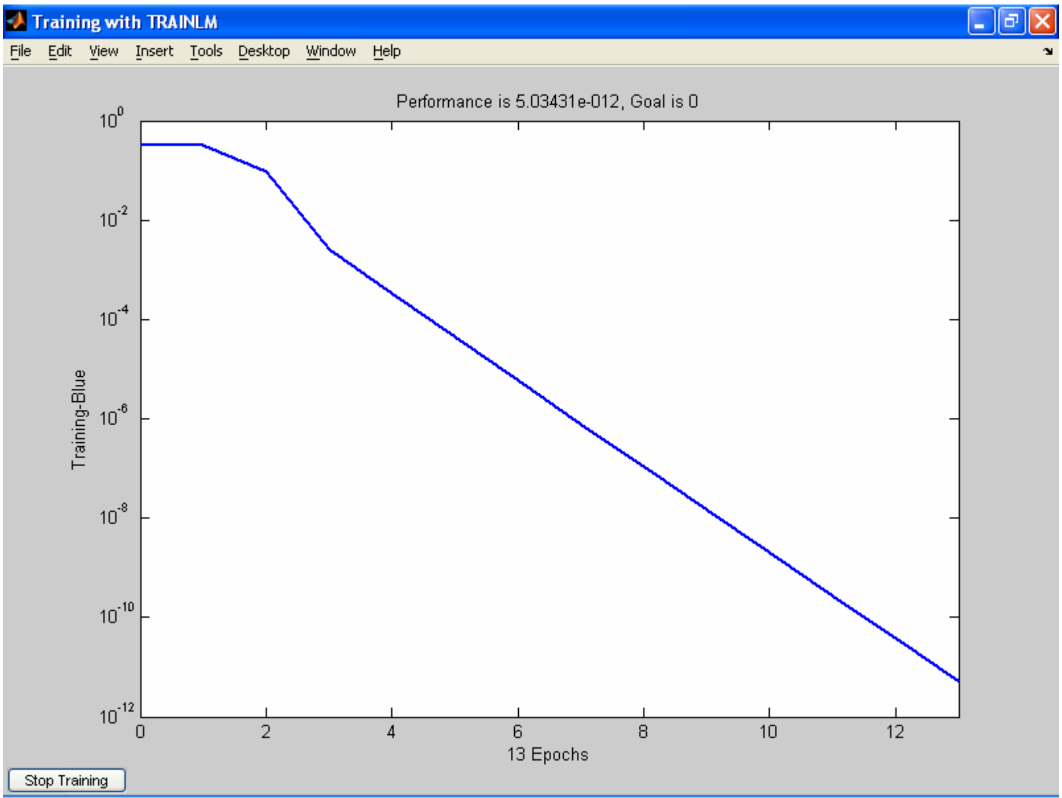
Table 2: Showing ratio values from subscriber data

Calling number	Calls made ratio	Calls received ratio	Calls dropped ratio	Amount charged ratio	Service utility ratio	Credit load ratio	ssb ratio	drc ratio	ilc ratio	imr ratio	rb ratio	gp ratio	wp ratio	vs ratio	dl ratio	hp ratio	vp ratio	Complaint frequency ratio	Response ratio	Mean-response	No-response ratio	Churn
080x1	17.00	13.72	0.29	429.88	2.36	270.41	0.15	0.05	0.38	0.24	0.03	0.07	0.09	0.13	0.01	0.07	0.06	2.88	0.96	4.96	0.04	1
080x2	19.23	14.56	0.31	414.29	5.00	312.09	0.09	0.23	0.11	0.11	0.12	0.06	0.13	0.06	0.02	0.06	0.01	3.85	0.86	3.34	0.14	1
080x3	25.60	1.60	0.30	890.00	3.68	68.00	0.20	0.05	0.15	0.08	0.04	0.04	0.01	0.25	0.10	0.05	0.04	0.65	0.97	14.29	0.11	0
080x4	13.70	5.51	0.05	342.48	2.70	328.77	0.18	0.08	0.05	0.12	0.09	0.05	0.08	0.12	0.05	0.01	0.18	2.60	1.00	2.06	0.11	0
080x5	16.89	14.81	0.04	210.71	1.25	187.28	0.08	0.09	0.06	0.00	0.23	0.05	0.07	0.08	0.09	0.18	0.06	2.04	0.82	10.44	0.18	1

Table 3: Showing rescaled values for neural network input (split into two halves)

Calling number	Calls made ratio	Calls received ratio	Calls dropped ratio	Amount charged ratio	Service utility ratio	Credit load ratio	ssb ratio	drc ratio	ilc ratio	imr ratio	rb ratio
080x1132110	0.019101	0.015410	0.000323	0.483011	0.002657	0.303833	0.000166	0.000060	0.000426	0.000268	0.000039
080x1332710	0.021608	0.016360	0.000346	0.465490	0.005618	0.350661	0.000104	0.000257	0.000128	0.000120	0.000133
080x5612770	0.028764	0.001798	0.000337	1.000000	0.004131	0.076404	0.000225	0.000056	0.000169	0.000084	0.000045
080x3434160	0.015392	0.006187	0.000062	0.384808	0.003038	0.369401	0.000198	0.000095	0.000059	0.000131	0.000106
080x1752819	0.018976	0.016646	0.000042	0.236754	0.001403	0.210424	0.000094	0.000103	0.000066	0.000005	0.000262

gp ratio	wp ratio	vs ratio	dl ratio	hp ratio	vp ratio	Complaint frequency ratio	Response ratio	Mean-response	No-response ratio	Churn
0.000074	0.000106	0.000147	0.000011	0.000080	0.000070	0.003232	0.001081	0.005568	0.000043	1
0.000072	0.000144	0.000064	0.000024	0.000063	0.000013	0.004322	0.000963	0.003757	0.000161	1
0.000039	0.000011	0.000281	0.000112	0.000056	0.000045	0.000730	0.001089	0.016051	0.000118	0
0.000053	0.000086	0.000132	0.000053	0.000012	0.000198	0.002924	0.001124	0.002318	0.000118	0
0.000059	0.000074	0.000089	0.000099	0.000201	0.000072	0.002289	0.000919	0.011730	0.000204	1



3.2 Survival Data mining

The survival time analysis of subscriber data to estimate the life time value of a subscriber and the churn hazard per subscriber over a time period will be carried out using SAS software [15],[16]. Survival data mining describes the distribution of the survival time for customers in a given population, investigates the strength of parameter influence on expected survival time and allows comparing of survival time distributions among different subpopulations [9]. This analysis gives insight into customers' behaviours and finds ways to increase their survival time. This phase of the experiment has not been concluded, it is intended to determine the survival rate of each subscriber in the population and predict at what point a subscriber is likely to churn.

3.3 Generating Reasons for Churn and Intervention Strategy

The results obtained from churn prediction using artificial neural network and churn survival analysis will be fed into a Decision Support Expert System (DSES) which will generate the probable reasons for churn and recommendations of intervention strategies for customer retention using a knowledge rule-based inference engine. The inference engine of the DSES is composed of a set of if - then rules that provides recommendations of appropriate incentives based on the credit rating of a subscriber. Subscribers are classified as high-valued, medium-valued and low-valued using a specific value rating threshold function Th(x) given as:

$$\mathbf{Th(x)} = \begin{cases} 1 & \text{if } \mathbf{V(x)} \geq 60 \quad \text{// 'high} \\ 0.5 & \text{if } \mathbf{V(x)} \geq 40 \text{ and } \mathbf{V(x)} < 60 \quad \text{// 'medium} \\ 0 & \text{if } \mathbf{V(x)} < 40 \text{ // 'low} \end{cases} \qquad (3)$$

Where:

V(x) = ((calls made ratio + calls received ratio) / 2 * 100) (4)

Generally, low valued customers will be ignored while medium and high valued customers will have best-fit retention strategies recommended for them using a set of in-built rules.

In other to generate reasons for churn, the DSES reckons a particular complaint ratio metric as 'high' if it is above 0.5 and 'low' otherwise. Typical instances of inference rules for generating reasons for churn are shown as follows:

1. *If call dropped ratio ≥ 0.5 ('high') and churn prediction = 1 ('Yes') Then 'subscriber may churn due to high dropped calls'*
2. *if ssb ratio ≥ 0.5 and call dropped ratio ≥ 0.5 and churn prediction = 1 Then ' subscriber may churn due to ssb problems, high dropped calls'*
3. *if ilc ratio ≥ 0.5 and ssb ratio ≥ 0.5 and call dropped ratio ≥ 0.5 and churn prediction = 1 Then 'subscriber may churn due to ilc, ssb problems, high dropped calls'*

4. Results

Thus far, we have completed the implementation of the neural network-based churn prediction and DSES components of the architectural framework, and the results obtained have been very encouraging. Although more time is still needed to conclude work on the aspect of survival analysis, our experience so far gives credence to the feasibility of this process framework as a viable model for effective subscriber management in the telecommunication industry. The DSES proves to be an ideal complement to the interpolative power of the neural network in predicting prospective future churn by providing the probable reasons why churn may occur and a suggestion of possible intervention strategies. All of this has so far been achieved with an acceptable level of accuracy.

5. Conclusion

A comprehensive process framework for management of churn that embraces churn prediction, determination of why and when of churn and automatic recommendation of retention strategies like the one described in this work offers a viable platform for effective subscriber management in any fast maturing telecom market. This kind of holistic approach that will ensure the retention of high-valued customers and ultimately the promotion of profitability is therefore recommended. We intend to complete the other aspects of this work in order to further drive home the novelty of our approach in contrast to many other existing approaches.

References

1. Au, W., Chan, C. C. and Yao, X. (2003): A Novel Evolutionary Data Mining Algorithm with Applications to Churn Prediction, IEEE transactions on evolutionary computation, 7, 6, 532-545.
2. Bloemer, J., Brijis, T., Vanhoof, K. and Swinnen, G. (2002): Comparing Complete and Partial Classification for Identifying Customers at Risk, International journal of research in marketing, 20, 117-131.
3. Boone, D. S. and Roehm, M. (2002): Retail Segmentation Using Artificial Neural Networks, International journal of research in marketing, 19, 287-301.
4. Daramola, J., Olugbara, O., Owoloko, E. and Musa, A (2006): Decision Support for Property Appraisal Using a Neural-CBR Intelligent System., Proceedings of the International Conference on New Trends in Mathematical and Computer Sciences With Applications to Real World Problems (NTMCS), Covenant University, Ota, Nigeria [65-82].

5. Euler, T. (2005): Churn Prediction in Telecommunication Using MiningMart, In: Proceedings of the Workshop on Data Mining and Business (DMBiz) at the 9th European Conference on Principles and Practice in Knowledge Discovery in Databases (PKDD), Porto, Portugal.
6. Hadden, J. Tiwari, A., Roy, R, and Ruta, D (2006): Churn Prediction using Complaints Data, Transactions on Engineering, Computing and Technology, Enformatika, Vol. 13, ISBN 1305-5315, pp 158-163.
7. Hwang, H., Taesoo, J. and Euiiho, S. (2004): An Ltv Model and Customer Segmentation Based on Customer Value: A Case Study on the Wireless Telecommunications Industry, Expert systems with applications, 26, 181-188.
8. Kim, H. and Yoon, C. (2004): Determinants of Subscriber Churn and Customer Loyalty in the Korean Mobile Telephony Market, Telecommunications Policy, 28, 751-765.
9. Linoff, G. (2004) : Survival Data Mining for Customer Insight, Data miners, http://www.data-miners.com
10. Lu, J (2001): Detecting Churn Triggers for Early Life customers in the Telecommunication Industry- An application of Interactive Tree Training, Proceedings of the 2nd Data mining Conference of DiaMondSUG 2001, Chicago IL.
11. Lu, J. (2002): Predicting Customer Churn in the Telecommunications Industry- An Application of Survival Analysis Using SAS, Proceedings of SUG127, Orlando, Florida, Paper 114-27.
12. Mathworks Inc. (2003): MATLAB © Software, version 7.0
13. Mozer, M., Wolniewicz, R., Grimes, D, Johnson, E., and Kaushansky, H (2000): Predicting Subscriber Dissatisfaction and Improving Retention in the Wireless Telecommunication Industry, IEEE Transactions on Neural Networks, Special Issue on Data Mining and knowledge Representation
14. Rygielski, C., Wang, J. and Yen, D. C. (2002): Data Mining Techniques for Customer Relationship Management, Technology in Society, 24, 483-502.
15. SAS Institute Inc. (1989), SAS/STAT @ Software: Changes and Enhancements through Release 6.11, Volume 2, Cary, NC: SAS Institute Inc. 846pp.
16. SAS Institute Inc. (1989), SAS/STAT @ Users Guide, Version 6.0, Fourth Edition, Volume 2, Cary, NC: SAS Institute Inc. 846pp.
17. Vellido, A., Lisboa, P. J. G. and Meehan, K. (1999): Segmentation of the Online Shopping Market Using Neural Networks, Expert systems with applications, 17, 303-314.
18. Wei, C. and Chiu, I. (2002): Turning Telecommunications Call Details to Churn Prediction: A Data Mining Approach, Expert systems with applications, 23, 103-112.

MOBILE INTERNET CRYPTOGRAPHY

T. OLORUNFEMI[1], B.K. ALESE[2] AND O. AKINSELE[3]

[1]*Computer Centre, Adekunle Ajasin University, Akungba-Akoko, Nigeria*

[2]*Computer Science Department, Federal University of Technology, Akure, Nigeria*

[3]*Computer Science Department, Nnamdi Azikiwe University, Awka, Nigeria*

Email: olorunfemitos@yahoo.com Tel:+234-803-675-0017 ; +234-805-853-1677

ABSTRACT

Electronic-tools for mobile internet such as pagers, personal digital assistants (PDAs), cellular phones and smart cards are often physically constrained. They also usually lack processing power, storage space, bandwidth, or power consumption hence the need to adopt the right cryptographic tool.
Several security tools such as Secure MIME, secure socket layer, Internet Security Protocol, secure shell, Kerberos and public cryptosystems were reviewed and the use of elliptic curve cryptosystem (ECC) was recommended for mobile internet because of its smaller key size, which leads to faster computation time, and reduction in processing power, storage space and bandwidth. This makes ECC ideal for constrained environments such as pagers, cellular phones and PDAs.

Keywords: cryptography, elliptic curve cryptosystems, and internet security.

1.0 Introduction: Most e-tools for mobile internet such as pagers, personal digital assistants (PDAs), cellular phones and smart cards are often physically constrained. They also usually lack processing power, storage space, bandwidth, or power consumption hence the need to adopt the right cryptographic tool to match the constrained environment can not be overemphasized.

2.0 Definition Of Cryptography: *Cryptography* is the study of mathematical techniques related to aspects of information security such as confidentiality, data integrity, entity authentication, and data origin authentication.(Menezes *et al., 1997)*. Cryptography is according to Aydos (2000) , in general, the science of concealing data. However, Alese (2004) defined it as the art and science of concealing data which is further asserted by Rabah (2004) and this same view is shared by Olorunfemi (2006).
Cryptology (from the Greek *kryptós lógos*, meaning "hidden word") is the study of codes. Cryptology is also the discipline of cryptography and cryptanalysis combined. The Jewish writers sometimes concealed the meaning of their writing by alphabet reversal algorithm by using the last letter of the alphabet in place of the first, the next last for the second, and so on. This system, called atbash, is exemplified in the Bible around 606BC, in Jeremiah 25:26, in which "Sheshach" is written for "Babel" (Babylon), using the second and twelfth letters from the end of the Hebrew alphabet instead of from the beginning (The Holy

Bible,2004). Around 50 BC, Julius Caesar developed the idea of transposing letters in the alphabet in order to transmit military messages with relative security (James, 2005).

Encryption is the transformation of data into a form that is as close to impossible to read without the appropriate knowledge. Its purpose is to ensure security by keeping information hidden from anyone for whom it is not intended, even those who have access to the encrypted data. Traditionally, secrecy has meant security, however Olorunfemi and Oladipo (2005) stated that this might be so in theory but not necessarily in practice, Olorunfemi and Akinsele (2006) also asserted this. *Decryption* is the reverse of encryption; it is the transformation of encrypted data back into an intelligible form.

Cryptography is fundamentally based on problems that are difficult to solve. A problem may be difficult because its solution requires some secret knowledge, such as decrypting an encrypted message or signing some digital document. The problem may also be hard because it is intrinsically difficult to complete, such as finding a message that produces a given hash value. (Rivest, 1990; Brassard, 1988; Stinson, 1995; and Stallings, 1995).

2.1 Public Key Cryptosystems: Diffie and Hellman introduced a key exchange protocol in their first publication along with their ideas of public key cryptography. Their protocol is known as *Diffie-Hellman key exchange*. Diffie and Hellman (1976a and 1976b) introduced the concept of a digital signature in 1976. Although the idea of a digital signature was clearly articulated, no practical realization emerged until the 1978 paper by Rivest, Shamir, and Adleman (Rivest *et al.,* 1978). Diffie and Hellman (1976a) suggested that public key schemes would ultimately be of more importance to the business community than the confidentiality services for which cryptography had traditionally been used. This research work is based on Elliptic Curve Cryptosystems.

Elliptic curve cryptosystems were first proposed independently by Victor Miller (Miller,1986) and Neal Koblitz (koblitz, 1987a) in the mid-1980s. Elliptic curve cryptography constitutes a fundamental and efficient technology for public key cryptosystems (Konstantinou *et al.,* 2003). At a high level, they are analogs of existing public-key cryptosystems in which modular arithmetic is replaced by operations defined over elliptic curves.

The security of many public-key cryptosystems relies on the apparent intractability of the computational problem they are based upon. In a cryptographic setting, it is prudent to assume that the adversary is very powerful. The digital signature schemes in use today can be classified according to the hard underlying mathematical problem, which provides the basis for their security:

2.1.1 Integer Factorization (IF) Schemes: They base their security on the intractability of the integer factorization problem. Examples include RSA (Rivest *et al.*, 1978) and Rabin (Rabin, 1979) Signature Schemes.

2.1.2 Discrete Logarithm (DL) Schemes: They base their security on the intractability of the (ordinary) discrete logarithm problem in a finite field. Examples of this include the ElGamal (ElGamal,1985), Schnorr (Schnorr, 1991), DSA (NIST,1994) and Nyberg-Rueppel (Nyberg and Rueppel, 1993 and 1996)

2.1.3 Elliptic Curve (EC) Schemes: They base their security on intractability of the elliptic curve discrete logarithm problem. An example is the Elliptic Curve Digital Signature Algorithm (Vanstone, 1992; Johnson and Menezes, 1999). The elliptic curve scheme is definitely the most recent of all the various schemes (Alese (2004), Rabah (2005a and 2005b)).

2.2 Applications of Cryptography : With the information-processing and telecommunications revolutions well underway, there is an increasing demand for techniques for keeping information secret, for determining that information has not been tampered with, and for determining who authored pieces of information. Cryptographic techniques are currently being utilized in the following application areas but the list is definitely not exhaustive:

- The Internet : Secure electronic mail ; Home banking ; Internet browsers
- The Financial Services Industry : Electronic cash ; Credit card transactions ; Instant teller banking ; Wholesale banking
- Wireless Communications : Pagers ; Cellular telephones ; Smart cards
- Telecommunications : Fax encryptors ; Modems ; Secure telephones ; Cable TV and pay-per-view

More applications of cryptography is available at www.cacr.math.uwaterloo.ca, www.rsasecurity.com, and Menezes *et al. (*1997)

Cryptography is not confined to the world of computers. Cryptography is also used in cellular (mobile) phones as a means of authentication; that is, it can be used to verify that a particular phone has the right to bill to a particular phone number. This prevents people from stealing (“cloning") cellular phone numbers and access codes. Another application is to protect phone calls from eavesdropping using voice encryption.

2.3 Cryptography and Internet: The internet in recent years has become a colossus over our everyday life so in order to appreciate cryptography better; we will quickly review its application in relation to internet security

2.3.1 (Secure / Multipurpose Internet Mail Extensions) – S/MIME: It is a protocol that adds digital signatures and encryption to Internet MIME (Multipurpose Internet Mail Extensions) messages. MIME is the official proposed standard format for extended Internet electronic mail. Internet e-mail messages consist of two parts, the header and the body. The purpose of S/MIME is to define such services, which itself is the result of cryptographic processing on other MIME body sections.

S/MIME has been endorsed by a number of leading networking and messaging vendors, including ConnectSoft, Frontier, FTP Software, Qualcomm, Microsoft, Lotus, Wollongong, Banyan, SecureWare, VeriSign, Netscape, and Novell. Information on MIME can be found at ftp://ftp.isi.edu/in-notes/rfc1521.txt.

2.3.2 Secure Socket Layer: The SSL (Secure Sockets Layer) Handshake Protocol (Hickman, 1996) was developed by Netscape Communications Corporation to provide security and privacy over the Internet. The protocol supports server and client authentication. The SSL protocol is application independent, allowing protocols like HTTP (HyperText Transfer Protocol), FTP (File Transfer Protocol), and Telnet to be layered on top of it transparently. Still, SSL is optimized for HTTP; for FTP, IPSec might be preferable. The SSL protocol is able to negotiate encryption keys as well as authenticate the server before data is exchanged by the higher-level application. The SSL protocol maintains the security and integrity of the transmission channel by using encryption, authentication and message authentication codes.

The SSL Handshake Protocol consists of two phases: server authentication and an optional client authentication. In the first phase, the server, in response to a client's request, sends its certificate and its cipher preferences. The client then generates a master key, which it encrypts with the server's public key, and transmits the encrypted master key to the server. The server recovers the master key and authenticates itself to the client by returning a message authenticated with the master key. Subsequent data is encrypted and authenticated with keys derived from this master key. A variety of cryptographic algorithms are supported by SSL. During the ``handshaking'' process, the RSA public-key cryptosystem is used. After the exchange of keys, a number of ciphers are used. More information on SSL 3.0, by Hickman (1996) is available at http://home.netscape.com/eng/ssl3/index.html.

2.3.3 Internet Security Protocols (IPSec): The Internet Engineering Task Force (IETF)'s IP Security Protocol (IPSec) working group is defining a set of specifications for cryptographically-based authentication, integrity, and confidentiality services at the IP datagram layer. IPSec is intended to be the future standard for secure communications on the Internet, but is already the de facto standard. The IPSec group's results comprise a basis for interoperable secured host-to-host pipes, encapsulated tunnels, and Virtual Private Networks (VPNs), thus providing protection for client protocols residing above the IP layer.

The protocol formats for IPSec's Authentication Header (AH) and IP Encapsulating Security Payload (ESP) are independent of the cryptographic algorithm, although certain algorithm sets are specified as mandatory for support in the interest of interoperability. Similarly, multiple algorithms are supported for key management purposes (establishing session keys for traffic protection), within IPSec's IKE framework.The home page of the working group is located at http://www.ietf.org/html.charters/ipsec-charter.html.

2.3.4 Secure Shell (SSH): Secure Shell, is a protocol which permits secure remote access over a network from one computer to another. SSH negotiates and establishes an encrypted connection between an SSH client and an SSH server, authenticating the client and server in any of a variety of ways (some of the possibilities for authentication are RSA, SecurID, and passwords). That connection can then be used for a variety of purposes, such as creating a secure remote login on the server (effectively replacing commands such as telnet, rlogin, and rsh) or setting up a VPN (Virtual Private Network).

When used for creating secure logins, SSH can be configured to forward X11 connections automatically over the encrypted “tunnel" so as to give the remote user secure access to the SSH server within a full-featured windowing environment. SSH connections and their forwarding can be cascaded to give an authenticated user convenient secure windowed access to a complete network of hosts. Other TCP/IP connections can also be tunneled through SSH to the server so that the remote user can have secure access to mail, the web, file sharing, FTP, and other services. The SSH protocol is currently being standardized in the IETF's SECSH working group: http://www.ietf.org/html.charters/secsh-charter.html. More information about SSH, including how to obtain commercial implementations, is available from SSH Communications Security (http://www.ssh.fi) and Van Dyke Technologies (http://www.vandyke.com).

2.3.5 Kerberos: Kohl *et al.,* 1994 stated that it is an authentication service developed by the Project Athena team at MIT. The first general use version was version 4. Version 5, which addressed certain shortfalls in version 4, was released in 1994. Kerberos uses secret-key ciphers for encryption and authentication. Version 4 could only use DES. Unlike a public-key authentication system, Kerberos does not produce digital. Instead Kerberos was designed to authenticate requests for network resources rather than to authenticate authorship of documents. Thus, Kerberos does not provide for future third-party verification of documents.

In a Kerberos system, there is a designated site on each network, called the Kerberos server, which performs centralized key management and administrative functions. The server maintains a database containing the secret keys of all users, authenticates the identities of users, and distributes session keys to users and servers who wish to authenticate one another. Kerberos requires trust in a third party (the Kerberos server). If the server is compromised, the integrity of the whole system is lost. Public-key cryptography was designed precisely to avoid the necessity to trust third parties with secrets. Kerberos is generally considered adequate within an administrative domain; however across domains the more robust functions and properties of public-key systems are often preferred. There has been some developmental work in incorporating public-key cryptography into Kerberos (Yaksha, 1995). Detailed information on Kerberos, is available at ftp://ftp.isi.edu/in-notes/rfc1510.txt.

3.0 Elliptic Curve Cryptosystem Implementation: Olorunfemi *et al*. 2006(a,b). implemented Elliptic Curve Cryptosystem (ECC) as an alternative to established public-key systems such as Digital Signature Algorithm (DSA) and Rivest Shamir Adleman (RSA). The main reason for the attractiveness of ECC is the fact that there is no sub-exponential algorithm known to solve the elliptic curve discrete logarithm problem on a properly chosen elliptic curve. Hence, it takes full exponential time to solve while the best algorithms known for solving the underlying integer factorization for RSA and discrete logarithm problem in DSA both take sub-exponential time.

This means that significantly smaller parameters could be used in ECC than in other competitive systems such as RSA and DSA, but with equivalent levels of security. A typical example of size in bits of the keys used in different public-key systems, with a comparable level of security (against know attacks), is that a 160-bit ECC key is equivalent to RSA and DSA with a modulus of 1024 bits (Olorunfemi *et al.* 2006c).

4.0 Conclusion: ECC provides more security per bit than RSA and DSA hence it uses smaller key. Some benefits of having smaller key sizes for mobile internet security include faster computation time, and reduction in processing power, storage space and bandwidth. This makes ECC ideal for constrained environments such as mobile phones, pagers, personal digital assistants (PDAs), cellular phones and smart cards. These advantages are especially important in other environments where processing power, storage space, bandwidth, or power consumption are lacking.

References

Alese, B.K. (2004), "Design of Public key cryptosystem using elliptic curve" Ph.D. Thesis , FUTA, Nigeria.

Aydos, M. (2000), "*Efficient Wireless Security Protocols based on Elliptic Curve Cryptography"*, Ph.D Thesis
Submitted to Oregon State University, 132pp

Brassard, G. (1988), *Modern Cryptology*, Springer-Verlag.

Diffie, W. and Hellman, M. E.(1976a), New directions in cryptography. *IEEE Transactions on Information Theory*,
22:644–654, November.

Diffie, W. and Hellman, M. E.(1976b), "Multiuser cryptographic techniques", *Proceedings of AFIPS National*
Computer Conference, 109–112.

ElGamal, T. (1985), "A public key cryptosystem and a signature scheme based on discrete logarithms", *IEEE*
Transactions on Information Theory, **31**, 469-472

Hickman, K.E.B. 1996. The SSL Protocol, available at http://www.netscape.com/eng/security/SSL_2.html.

James, H. (2005), Available at http://www.james-huff.com/math/caesar-history.htm

Johnson, D. and Menezes A. (1999), The elliptic Curve Digital Signature Algorithm (EDSA) Technical Report
CORR 99-34. Available at http://www.cacr.math.uwaterloo.ca

Koblitz, N. (1987), Elliptic curve cryptosystems, *Mathematics of Computation* **48** , 203-209.

Kostantinou E., Stamatiou, Y.C.,andZaroliagis C. (2003), "On the Efficient Generation of Elliptic Curves Over
Prime Fields" in CHES 2002, LNCS 2523, Springer-Verlag, Berlin. pp.333 – 348

Kohl,J., Neuman, B. and Tso, T. (1994), The evolution of the Kerberos authentication service, Distributed Open
Systems, *IEEE Press*.

Menezes, A. (1995), Elliptic Curve Cryptosystems, *CryptoBytes* (2) **1** (Summer)

Menezes, A. , Van Oorschot, P. and Vanstone, S. (1997), *Handbook of Applied Cryptology"* CRC Press available at
http://www.cacr.math.uwaterloo.ca

Miller, V.S. (1986), Use of elliptic curves in cryptography, *Advances in Cryptology - Crypto '85*, Springer-Verlag ,
417-426.

Multipurpose Internet Mail Extension available at ftp://ftp.isi.edu/in-notes/rfc1521.txt

National Institute of Standards and Technology .(1994), *Digital Signature Standard,* FIPS Publication 186, 1994.
Available at http://csrc.nist.gov/fips

Nyberg, K. and Rueppel, R. (1993), "A new signature scheme based on DSA giving message Recovery" *1st ACM*
Conference on Computer and Communications Security, 58-61

Nyberg, K. and Rueppel, R. (1996), "Message recovery for signature schemes based on discrete logarithm problem"
Designs, Codes and Cryptography, **7**, 61-81

Olorunfemi, T. and Oladipo, O.F. (2005), Security and open source software: A critical analysis. Conference

Proceedings (ed. Itegboje A.O.). 8th International Conference: *Information Technology Capacity Building: The future of Nigeria's economic growth*. **Vol. 16:** 277-282

Olorunfemi, T. (2006), "Design and implementation of elliptic curve digital signature algorithms" M.Tech Thesis, FUTA, Nigeria, 106pp.

Olorunfemi, T.and Akinsele, O. (2006). Critical Appraisal of Open Source Software Security. *Trends in Applied Sciences Research*. **Vol. 1(6):** 597-604

Olorunfemi, T.O. S. ; Alese, B.K.; Falaki, S.O. and Fajuyigbe, O. (2006a).Implementation of Elliptic Curve Digital Signature Algorithms - 1st International Conference on "New Trends in the Mathematical & Computer Sciences with Applications to Real World Problems. (Ed. C. K. Ayo *et al.*) *Conference Proceedings* , pp 523 – 544.

Olorunfemi, T.O. S. ; Alese, B.K. and Falaki, S.O.(2006b). Implementation of Elliptic Curve-Based Nyberg-Rueppel Digital Signature Algorithm. *International Journal of Numerical Mathematics.* **Vol. 2:** pp323-349

Olorunfemi, T.O. S. ; Alese, B.K.; Falaki, S.O. and Fajuyigbe, O. (2006c). Comparative Analysis of Elliptic Curve Cryptosystems with Conventional Cryptosystems. *Science Research Annals*, **Vol. 2(1)**: pp 19-27

Rabah, K. (2004), Data Security and cryptographic techniques – A review. *Info. Tech. J.* **3(1):** 106-132

Rabah, K. (2005a), Theory and Implementation of elliptic curve cryptography. *J. Applied Sci.,* ***5*** *:*604-633

Rabah, K. (2005b), Elliptic curve ElGamal encryption and signature schemes. *Info. Tech. J.* **4(3):** 299-306

Rivest, R.L.(1990), Cryptography, *Handbook of Theoretical Computer Science, volume A* (editor: J. van Leeuwen), MIT Press/Elsevier, Amsterdam, 719-755.

Rivest, R., Shamir, A. and Adleman L. (1978), "A method for obtaining digital signatures and public key cryptosystems" *Communications of ACM*, **21**, 120 -126

RSA Security available at www.rsasecurity.com

Schnorr, C. (1991), "Efficient signature generation by smart cards", *Journal of Cryptology*, **4** , 161-174

Secure Shell available at http://www.ietf.org/html.charters/secsh-charter.html.

Secure Socket Layer available at http://home.netscape.com/eng/ssl3/index.html

SSH Communications Security available at http://www.ssh.fi/

Stallings, W. (1995), *Network and Internetwork Security - Principles and Practice*, Prentice-Hall, New Jersey.

Stinson, D. R. (1995), *Cryptography - Theory and Practice*, CRC Press, Boca Raton.

The Holy Bible (2004) Authorized King James Version, Bible Society of Nigeria.

The Kerberos Network Authentication Service (V5) (J. Kohl and C. Neuman, RFC 1510) at ftp://ftp.isi.edu/in-notes/rfc1510.txt.

Van Dyke Technologies (http://www.vandyke.com).

Vanstone, S. (1992), "Responses to NIST'S Proposal", *Communications of ACM,* **35,** 50 – 52

Yaksha , R. Ganesan. (1995), Augmenting Kerberos with public key cryptography, *Proceedings of the 1995 Internet Society Symposium on Network and Distributed Systems Security*, IEEE Press , 132-143.

MOBILE PHONE: THE PAST, THE PRESENT AND THE FUTURE

M.O. OYELAMI AND A.A. AZETA

Computer and Information Sciences Department

Covenant University

Ota, Ogun State

Nigeria

olufemioyelami@yahoo.com, Azeta_ambrose@yahoo.com

Abstract

Since the introduction of cellular analogue phone in 1985 (first generation mobile phone), there has been a continuous improvement from the first generation to digital second-generation to 2.5 generation and now the third-generation. The ground is also being prepared for the fourth-generation mobile phone. Mobile technology has transformed our lives in ways that might have seemed unimaginable some years ago and yet we are still witnessing more transformations and many more are yet to come. This paper examines the concept of cellular communication, the development of mobile phones, the features in the past, the current trends and what to the future holds in general and specifically for Nigeria. The cost implications of the various generations over the previous ones to the end users are also discussed.

1. Introduction

Generations in the development of mobile telephone networks actually means improvement on services: new service providing logic and new telecommunication network architecture. First-generation cellular system was an analogue Frequency Modulation (FM) system which was entirely for voice transfer [1] and was introduced in 1985[3] but had capacity limitations of Time Division Multiple Access (TDMA), incompatibility problem, coverage limited to a geographical entity, low speech quality, no open interfaces except the radio interface and lack of security in speech transmission[2]. The second-generation mobile cellular systems use digital radio transmission. The boundary between the first- and the second-generation systems is the analogue/digital split. The second-generation networks have much higher capacity than the first-generation systems. One frequency channel is divided simultaneously among several users either by Code Division Multiple Access (CDMA) or time (TDMA) division [1]. They are structured atop existing first-generation analogue technology and are premised on compatibility and parallel operation with analogue networks [2].

Generation 2.5 is a designation that broadly includes all advances upgrades for the second-generation networks. The third-generation systems can be used not only for person-to-person communication but also for person-to-machine communication based on Wideband Code Division Multiple Access (W-CDMA) radio technology [12]. The 3G systems are targeted to offer a wide variety of services such as telephony, teleconference, voice mail, message broadcast navigation, etc [1] and will pave the way for 4G, which is likely to be the first truly converged network [10] which will incorporate new services more advanced than what the 3G offers.

1.1 Concept of Cellular Communication

The cellular concept uses a large number of low-power transmitters designed to serve only a small area. Instead of using a large transmitter to cover a wide area, the area is divided into smaller areas called cells. By reducing the total coverage area into small cells it is possible to reuse the same frequency in different cells without disturbing interference. The essential elements of a cellular system are:

1. Low-power transmitter and small coverage areas or cells;
2. Frequency reuse;
3. Handoff and central control;
4. Cell splitting to increase call capacity [1].

2. Overview of Network Generations

2.1 First-Generation Networks

The first-generation cellular systems were analogue FM systems, which were designed only for voice transfer. There was not a dominant standard but different competing ones. Among the successful ones were Nordic Mobile Telephone (NMT), Total Access Communications System (TACS) and Advanced Mobile Phone Service (AMPS)[1,2]. Other standards were often developed and used in one country, for instance C-Nerz in West Germany and Radiocomm 200 in France. In the UK (TACS), it operates in the 90 MHz waveband providing 1000 duplex channels occupying the frequencies 890 – 915 MHz (25 MHz) and 935 – 960 (25 MHz). In the US (AMPS), it operates in the 832 MHz waveband providing 1000 duplex channels occupying the frequencies 825 – 845 MHz (20 MHz) and 870 – 890 (20 MHz). The lower frequency band is assigned for transmission from mobile to mobile [1]. The first-generation networks were dated back to 1985[3].

2.1.2 Access Method

The first-generation mobile networks used Frequency Division Multiple Access (FDMA) technology. With FDMA, each service provider divides is spectrum of radio frequencies into individual frequency channels. Each channel has a width of 10 to 30 kilohertz (KHz) and is a specific frequency that supports a one-way communication session. For a regular two-way phone conversation, every cell phone caller is assigned two frequency channel one to send and one to receive [13]. Problems of the first-generation networks are:

1. They have capacity limitations of Time Division Multiple Access (TDMA) which divides the radio spectrum into multiple 30-KHz frequency channels;
2. They are incompatible systems. For instance, the US, Japan and the European Union have different standards;
3. They have their coverage limited to national boundaries;
4. They do not have interfaces except the radio interface;
5. They have low speech quality and
6. Lack of security in speech transmission.

2.2 Second-Generation Networks

The second-generation systems dated back to 1991 are digital and are capable of providing voice/data/fax transfer as well as a range of other value-added services. They have much higher capacity than the first-generation systems. One frequency channel is simultaneously divided among several users either by code (CDMA) or time (TDMA). There are four main standards for second-generation systems: Global System for Mobile (GSM) Communication and its derivative, Digital Advanced Mobile Phone Service (D-AMPS), Code Division Multiple Access (CDMA [IS-95]) and Personal Digital Cellular (PDC). GSM is by far the most successful and widely used 2G system [1]. The basic GSM uses the 900 MHz band, but there are also two derivatives: Digital Cellular System 1800 MHz and

PCS 1900. The reason for the new frequency band was the lack of capacity at 900 MHz band. The 1800 MHz band can accommodate far greater user population and thus has become quite popular especially in densely populated areas. The coverage area of the 1800 MHz is smaller than 900 MHz networks, and thus dual-band mobiles are used, where the phone uses a 1800 MHz network when such is available and otherwise roams to a 900 MHz network. D-AMPS is backward compatible with AMPS. It uses a digital control channel whereas AMPS has an analogue control channel [1]. Digital cordless systems are also mentioned as part of the second-generation systems. Two examples of these are: Digital Enhanced Cordless Telecommunications (DECT) and Personal Handyphone System (PHS). They do not have a network component. A typical system configuration includes a base station and a group of handsets. The base station is attached to some other network, which can be either a fixed or mobile network. The coverage area is often quite limited, consisting of town centres or office buildings. PHS is an advancement system and can do many things usually associated with mobile cellular systems [1].

2.2.1 Access Method

The second generation networks make use of CDMA or TDMA. Code Division Multiple Access (CDMA) instead of dividing a spectrum of radio frequencies into narrow frequency bands or time slots, uses a very large portion of that radio spectrum called frequency channel. Frequency channel has a wide width of 1.25 mega hertz (MHz) and is segmented into 64 multiple channels using a code to identify users. For duplex communications, each cell phone uses two of these wide CDMA frequency channels: one to send and the other to receive. CDMA is more immune to interference than TDMA and supports more users. Its capacity is not fixed but depends on coverage and total voice bandwidth [3].

Time Division Multiple Access (TDMA) divides the radio spectrum into multiple 30-KHz frequency channels. Every two-way communication requires two of these frequency channels: one to receive and the other to send. TDMA further subdivides each frequency channel into three to six time slots called voice/data channels, so that up to six digital voice or data sessions can take place using the same frequency [13]. The capacity is dependent on the number of available timeslots and cell phones for TDMA networks will not work with CDMA networks and vice versa[3]. Problems of the second-generation network are:

1. The biggest problem with the plain GSM is its low interface data rates;
2. The 2G network uses either CDMA or TDMA but cell phones for TDMA networks will not work with CDMA and vice versa;
3. For CDMA, cell size reduces due to spectrum use during peak periods;
4. For TDMA, capacity is dependent on the number of available timeslots.

2.3 2.5 Generation Network

This improves on the second-generation and there are three types of improvements:

a) **HSCSD (High Speed Circuit Switched Data):** It is the easiest way to speed up things. With HSCSD a mobile station can use several time slots for data connection. Currently, commercial implementations have a maximum of four time slots. The total rate is the product of the time slots and the rate of one slot. Its merits and demerits are those of circuit switching technique because it is circuit switched [1].
b) **GPRS (General Packet Radio Services):** This enables data rates to be pushed up to 115kbs and it is packet switched and as a result does not allocate the radio resources continuously but only when there is something to be sent. GPRS is suitable for non-real-time applications like e-mail and the Web surfing. It is not suited for real-time applications because the resource allocation in GPRS is connection based and thus it cannot guarantee an absolute maximum delay [1]

c) **EDGE (Enhanced Data Rate for Global Evolution):** It uses a modulation scheme called eight-phase shift keying (8PSK). This increases data rates of standard GSM networks by up to threefold. It requires a software upgrade to base station if the Radio Frequency (RF) amplifiers can handle high non-constant envelope modulation with EDGE's relatively high peak-to-average power ratio.

2.4 Overview of the Third Generation Systems

The year 2003 saw the full fledged appearance of what have been called third-generation mobile phones [4]. Third-generation systems are targeted to offer wide variety of services such as telephony, teleconference, voice mail, message broadcast navigation, location information, etc. Most of the services are wireless extension of Integrated Services Digital Network (ISDN) whereas services such as navigation and location information are mobile specific. Wireless network users will expect a quality of services similar to that provided by the wireline networks such as ISDN. Third-generation network will concentrate on the service quality, system capacity and personal and terminal mobility issues. The systems will be improved by using smaller cells and the reuse of frequency channels in a geographically ordered fashion [1]. Applying high-speed data transfer and state-of-the-art radio terminal technology, third-generation systems enable multimedia. It will be a catalyst for a whole new set of mobile service, enabling us to access advanced services anytime. It will free us from the confines of cables, fixed access points and low connection speeds. 3G brings together high-speed radio access and Internet Protocol (IP)-based services into one, powerful environment. IP is packet based, which means that users can be "online" at all times, but without having to pay until data is actually received or sent [2].

2.5 Data Transfer Rates of the Different Generations of Mobile Phone

The following are the data transmission rates of the different generations of mobile phone:
First-Generation: NMT- 380 bits per second (bps), 600 – 1,200 bps [19], TACS – 8 kilobits per second (kbps), AMPS – 10kbps [17].
Second-Generation: CDMA – 307kbps [20], TDMA – 10kbps [22].
2.5-Generation: GPRS – 115 kbps, EDGE – 384kbps [23], HSCSD – 43.2kbps [24].
Third-Generation: For high mobility – 144 kbps, for restricted mobility – 384 kbps and for indoor office environment – 2 mbps [26].

2.6 Cost Implications of the Different Generations of Mobile Phone to the End Users

The second-generation phones provided better quality and higher capacity at lower cost to end users [18]. GPRS which belongs to the 2.5 generation brings cost savings to both the mobile operator and consumers because GPRS radio resources are only needed while transferring the message. For the end user, it means you only pay for the time it takes to download [20].

High investments running into hundreds of billions are required in the third-generation mobile communication in the face of highly uncertain demand for 3G services and products [27]. The five UK licenses for the third-generation frequencies have fetched over twenty billion pound (£20,000,000,000) [6] and what this means is that the huge price paid for the licenses suggests that the facilities will not come cheap to the users at take-off. But, allowing more operators to come in will eventually force the price down as witnessed in Nigeria when Globacom rolled out its services. Generally, improvements from one generation to another enable reduced cost for the end users.

3. What the Future Looks Like

In 2003, full-fledged appearance of third-generation mobile phones was made. Data transfer rate is more than 100 times faster than that of second-generation mobile phones. This makes it possible to send and receive TV-quality images. There are now even some mobile phones that can receive actual TV broadcasts. And because a high-speed internet connection is possible, the fusion of mobile phones with personal computers is moving ahead [4]. The future of mobile phone is hard to predict, and trying is asking to be proved wrong. Nonetheless, the following are to be watched out for in the future and in the areas of mobile phone application:

i. **Fourth-Generation System:** Research into 4G system will be intensified and the focus will be in the areas of improving data transfer, capacity, introduction of diverse services and network transport protocols. The 4G system is likely to be the first truly converged network [10].

ii. **Advertisement:** Technological trajectories will likely continue for many years, making the phone a portable entertainment player, a new marketing tool for retailers and manufacturing, a multi-channel shopping device, a navigation tool, a new type of ticket and money and a new mobile Intranet device [11].

iii. **Larger Display:** Japanese firms have increased the size of phone display from a maximum of 2.0 inches (on the diagonal) in 1999 to 2.4 inches in 2003. New technologies will likely continue in this trend [11].

iv. **Processing and Network Speeds:** More immediate effects are expected from increases in processing power, memory and network speeds since they can improve the user interface without increases in the size of the display. Increased processing and memory capabilities reflect Moor's law. Decreasing semiconductor line widths have caused computing speed and memory size to roughly double every 18 months for the last 42 years. Similar trends are also seen in the mobile Internet, where the need for lower power consumption requires different circuit designs. Phones released in 2003 had speeds in the 100MHz to 200 MHz range and speeds greater than 500 MHz is expected in the future. Phones with more than 5 megabytes of internal memory were also released; some could save 2,000 photos (taken with a 300,000 pixel camera), 2,000 ring tones (with 40 polyphonic tones) or 100 Java programs. Phones with higher capacities than these are expected. Networks speeds will increase through the diffusion of third-generation services and other forms of networks are also expected to play an important role in the mobile internet [11].

v. **Growth, New Services and Technologies:** Since 2001 when the first GSM operator in Nigeria was licensed [5], the number of subscribers to the different GSM networks in Nigeria has continued to skyrocket. There were 113.55 million African mobile subscribers by the end of 2005 and it is forecast that it will rise to 378 million by 2011 [8]. As price competition among the various GSM network operators forces margins down, the mobile networks will move to offer additional services to increase revenue and customer commitment [6].Mobile phones are evolving very fast. Today's latest handset model is new for just a few weeks, and obsolete in a few months. The network technologies used are moving on fast. It is expected that new technologies mobile phones will be the order of the days ahead. Many people expect that phones will continue to become smaller and predict the arrival of models that can be worn as a wristwatch or pendant [6].

vi. **CD-Quality Music Will Be The Killer Cell Phone Application:** Within four years compact disk (CD)-quality music is going to be on more than half the entry-level cell phones used around the world. Music will be the killer application on cell phones measured by consumer use and revenue-generated by handset makers and wireless service providers [7].

vii. Mobile TV Use on Cell Phones Will Grow, but Not as Fast as CD-Quality Music:

Mobile TV will continue to be developed in the next few years, but will be challenged to win market acceptance because of the small size of cell phone screens. Music conversely, will be a non-stop, slam dunk appealing feature of a cell phone that won't require staring into a small screen [7].

viii. Digital Cameras on Cell Phones Will Have Less Customer Adoption Compared with Music: A cell phone camera generally produces average-to-poor quality photos compared with stand-alone cameras, which is no wonder when looking at the price. Average cell phone cameras cost virtually nothing whereas stand-alone cameras have a price tag similar to a feature phone. Further still, only a small percentage of cell phone users, similar to the overall population at large, are avid users of cameras. This percentage is much lower than the number of people who like to listen to music. Cameras will not disappear from cell phones, but usage of average cell phone cameras will be low compared to music players [7].

ix. Simultaneous Cell Phone Applications Will Be Crucial: Cell phones that can do most simultaneous applications like listening to music when receiving and answering a text message and playing music while a call comes in, downloading a video clip while listening to music etc at the most affordable costs will win over customers[11].

x. More Feature-Rich Cell Phones Will Be Key in Enabling Wireless Service Providers to Boost Their Average Revenues Per User: Wireless service providers will continue to grapple with a problem that has been vexing them for years. The problem is that even though the growth rate for cell phone sales and cell phone subscribers continues to increase, the average revenue the service providers and handset manufacture gain from each subscriber continues to fall. To increase that revenue, a growing number of service providers will continue to pitch and offer cell phones and subscribers on more feature-rich, revenue-generating services such as high quality audio streaming, download portals for music and video, and mobile TV [7].

xi. 3G Cell Phones Will Not Be All About Merely Enabling the Connected Lifestyle, but Rather Perfecting the Connected Lifestyle: The cell phone will be central for perfecting the connected lifestyle. This means that in the future, reliable, consistent, always-on connectivity will become essential anywhere, anytime. It will become less acceptable during the next few years to just make a cell phone connection. The connection will need to be perfected, meaning not flawed and unreliable, but rather as good as it can be. Lousy service, interrupted calls, the inability to connect and stay connected will become less tolerated than ever. An industry-changing mindset will take over, in which it won't be about being wirelessly connected. It will be about being wirelessly connected quickly, easily, inexpensively and reliability-in other words, perfectly [7].

xii. More Collaborations Will Be Necessary to Survive and Thrive in the Cell Phone Market:

Wireless service providers are teaming with the holders of music and video rights to offer downloads, and they will be building relationships with broadcasters to deliver mobile TV without burdening the mobile network. On the handset side, a rapidly growing number of collaborations will be formed between hardware and software companies to build more complete, more reliable, and more feature-rich cell phone platforms. Fewer companies will be able to do everything required to build a platform on their own. Cellular baseband experts, for example, will collaborate with providers of global positioning system solutions, for installation in these basebands, to provide value-added functionality. These collaborations will help accelerate time to market while minimizing research and development expenses.

xiii. More Use Will Be Made of Mobile Phones in the Education Sector: In the education sector, mobile phones are going to be used more extensively for learning purpose. Some education authorities are already sending truant 'text alerts' to parents. They are also being used for notification of lesson and classroom changes and cancellation of lectures and to inform students that library books are overdue. When camera phones are more commonplace, students could make notes and take photographs on field trips and file their reports using those pictures and notes. The notes could be on mobile device which is synchronized with a PC or server for immediate processing and results [13].

xiv. Convergence: In the short-to mid-term, the mobile industry will most likely follow a paradise lost scenario, in which existing players will find their market share and their profits threatened by new entrants. Cooperation among industry players is likely. Alliances to save costs and to pool parts of the value chain are already occurring, for example between Sharp and Sony Ericcson for 3G Research and Development activities. Personal Digital Assistants (PDAs) and mobile phones will no longer be distinct devices and increasing number of handsets will continue to incorporate digital cameras as a standard feature [10].

4. Nigeria and the Future of Mobile Phone

The growth of mobile phone is greater in Nigeria than in any other country in the continent [15] and as such, the future looks promising much more therein than in any other place. As at April, 2002, Mobile Telephone Network (MTN) and Econet provided six hundred thousand (600,000) lines [15] and as at December, 2005, nineteen million (19,000,000) were provided by all the Global System for Mobile Communications (GSM) service providers. This number is expected to at least double in the next five years. The expected increase in the number of mobile phones users in Nigeria will culminate in more foreign investment especially in the area of handset manufacturing thereby creating more employment opportunities and boosting the economy. Few banks in Nigeria like Guarantee Trust bank (GTB) already offer mobile banking services. With the expected rise in the number of mobile phones, the expectation is that more banks will join the league of those already offering the service. This of course, will boost the income generation capacity of the different mobile networks on whose platforms the service is enabled. More corporate organizations will make use of mobile phone for advertisement as currently being done by Spring bank in connection with Globacom. There will also be a proliferation of mobile banking applications as more and more organizations register their presence in the mobile world. Tariffs competition among the various GSM networks will eventually force the price down as currently being witnessed between Globacom and MTN. Universities like Covenant University and University of Ibadan, have already introduced mobile programming curricula. More schools are expected to follow suit.

5. Conclusion

The introduction of the first-generation analogue mobile phones brought transformation into our lives though not without their limitations. The second-generation systems brought in much more transformation and high capacity services. The concentration of the third-generation network will be value-added services, service quality, system capacity and personal and terminal mobility issues. As the third-generation mobiles prepare ground for the fourth-generation, a range of new services, technologies and applications will be introduced. Convergence in terms of alliances on the part of existing players in the industry will not be inevitable to save cost, as each one will find its market share and profit being threatened by the new entrants. Nigeria as a nation with the highest potential growth for mobile phone in Africa will not be left out of reaping the benefits of the future especially economic wise.

References

[1] Advances in Wireless Concept of Cellular Communication, http://cbdd.wsu.edu/kewlcontent/cdoutput/TR502/page69.htm

[2] Telecommunications architectures, http://www.tml.tkk.f1/studies/TIK-110.300/2000/Homeworks/assignment_14_model.html

[3] Othman Laraki, 15.020 Competition in Telecoms Recitation#4, http://ocw.m.t.edu/NR/rdonlyres/sloan-School-of-Management/15-020fall2003/A4A/AF2E-1795-4B94-Ac15-A39133325988C/drecl.pdf

[4] The future of Mobile Phones, http://web-japan.org/kidsweb/techno/mobile/future.html

[5] Aliga, P.A, Onianwa, C.U., Sadiq, I.F. 2003, GSM: The Economic Implications in Nigeria. Proceeding of the Nigeria Computer Society, vol 14 No. 1 page 187.

[6] History: Future, http://www.mobileshop.org/history/future.htm

[7] Predictions: Agere Systems 10 predictions about the cell phone industry in 2007, http://www.theunwired.net/?itemid=3478

[8] Mobile industry key Africa'future, http://www.hysorg.com/newso545.htm/#

[9] NathanEagle:Kenya and the future of mobile phone, September 11, 2006. http://www.boingboing.net/2006/09/11/nathan_eagle_kenya_a.html

[10] Convergence is now: Accenture Mobile Handset Study, http://www.cg163.net/sorthtm/d.aspid=34633.htm

[11] Microwave in Europe, http://www.findarticles.com/p/articles/mi_moNTN/is_50/ai_1115068875

[12] O.B. Okunoye, A.O. Adewumi, C.O. Uwadia, 2003: On the Evolution of Next-Generation GSM. Proceedings of the Nigeria Computer Society, vol 14 No. 1, pages 207-218

[13] viewpoint_vision article, The future of mobile technology; learning ‘on the run’ http://futurelab.org.uk/viewpoint/vision/vision_01__04.htm

[14] Held Gilbert , 2003, Building a Wireless Office, Auerbach Publications.

[15] Dan Isaacs (2002). Nigeria goes mad for mobiles, http://news.bbc.co.uk/1/hi/business/1905344.stm

[16] Wikipedia, free encyclopedia, Communications in Nigeria, http://en/wikipedia.org/wiki/communications_in_Nigeria

[17} C.K. Toh (2002). The Wireless Network Evolution. http://www.phptr.com/articles/article.asp?p=26330&rl=1

[18] Evolution of the Mobile Technology, http://www.buzzle.com/editorials/7-18-2004-56792.asp

[19] Nordic Mobile Telephone, http://www.answers.com/topic/nordic-mobile-telephone

[20] PCTechGuide, 2003, http://www.pctechguide.com/74MobileComms_GPRS.htm

[21] History of mobile phones, From Wikipedia, http://en.wikipedia.org/wiki/History_of_mobile_phones#First_Generation_Cellular

[22] C.K. Ayo, Phd lecture notes on Computer Communications Networks and Distributed Processing, 2006 (unpublished).
[23] Mobile Networking: 1G to 4G, http://www.ja.net/Services/publications/factsheets/057-mobile-networking.pdf
[24] Lionel M. Ni, Pei Zheng (2005), Smart Phone and Next Generation, Morgan Kaufmann
[25] Wikipedia, Encyvclopedia http://www.en.wikipedia.org/wiki/4G
[26] Jerry D. Gibson (1999), The Mobile Communications Handbook, CRC Press LLC.
[27] Martin Fransman, The impact of the 3G Mobile License Aunctions http://www.teleomvisions.com/current/auction.php

ONTOLOGIES AND SEMANTIC WEB SERVICES FOR M-COMMERCE APPLICATIONS

E.E.WILLIAMS [1], A.E. IBOR [2] AND U.G. INYANG [2]

[1]Department of Computer Science, University of Calabar, Calabar.

[2]Department of Mathematics, Statistics & Computer Science, University of Uyo, Uyo.

e-mail: udoiinyang@yahoo.com

Abstract

The Semantic Web provides the infrastructure that enables the semantic interoperability of web services. Semantic Web Services, which are services that are self-described and amendable to automated discovery, composition and invocation, will form the next generation of the Web. These services can be represented using ontology and rendered for use in e-commerce vis-a-vis m-commerce. This paper gives a design based on ontologies and semantic web services for m-commerce applications in the domain of Internet bookshelf. This design enhances the security, communication and interoperability between systems, customers and objects and provides a framework for effective m-commerce applications.

Keywords: Ontology, Semantic Web, Semantic Web Services, e-commerce, m-commerce, online bookstore

1. Introduction

We define electronic commerce (e- commerce) as business processes on the Internet. For example, advertisements; buying and selling of goods and services on the Internet. These transactions can be carried out between either two or more business operators (B2B) or- between business operators and individual and consumers (B2C). Among the numerous factors that contribute to the development of e-commerce are the tremendous development of the Internet and related technologies, the understanding and exploitation of the business potentials that rest behind this development. The relative ease and varieties consumers have when buying and selling on the Internet have led to exponential rise in e-commerce globally.

The facilities in e-commerce extend to include the devices that the end-user employs to gain access to the Internet. The main technologies developed to gain access to the Internet are the wired- line (e.g. using a home PC as end user device) or wireless (e.g. using a mobile phone as end user device). While e-commerce continues to impact the global business environment profoundly, technologies and applications are beginning to focus on mobile computing and the wireless web [Tarasewich et al, 2002]. Mobile Commerce (also called m-commerce, mobile e-commerce) involves all activities related to commercial transactions conducted through communications network that interface with wireless (or mobile) devices [Tarasewich et al, 2002]. That is, m-commerce is concerned with the use of wireless devices to gain access to the Internet for the purpose of carrying out business processes (or e-commerce). M-commerce therefore is a subset of e-commerce [Knospe and Schwiderski-Grosche,2002]. The main driving force for the rapid acceptance of mobile devices (for example bluetooth, smart phones, Mobile phones, laptops, earpiece, GSM etc) is the capability to get services and run applications at any time and at any place, especially while on the move [Veijalanen et al , 2003;Gray, 2003].

Semantic Web is a web of meaningful contents and services, which can be interpreted by computer programs. Semantic web addresses the weaknesses of current knowledge management systems; such as extraction of un-integrated relevant information from different sources, inability to maintain structured sources of information leading to difficult updating processes, unavailability of machine processable semantic representation of information to enhance automation, etc. On the other hand, Ontologies provide a shared and common understanding of a domain that can be communicated between people and application systems; a way to cope with heterogeneous representations of web resources. [Veijalainen et al, 2003] proposes an ontological transaction monitor which acts as an intelligent component between the application and the servers accessed during m-commerce transactions; and controls the perceivable communication behaviour of the terminal towards the servers. The main issues and activities in m-commerce are discussed in [Tarasewich et al, 2002].

There are demands for efficient ways of processing data especially on the web. The tools to use become even more important as the society becomes more complex with associated large volumes of data. M-commerce has additional advantages of convenience; accessibility and localization hence used by a lot more people, thereby increasing the volume of data and complexity of transactions on the web. Ontology and semantic web services are the tools suitable for processing this type of complex, voluminous amount of data, and enhances interoperability and communication between the heterogeneous systems [Williams et al, 2004]. These indicators form the driving force in carrying out this study.
In this paper, a design of ontology based semantic web services for effective m-commerce transactions in the domain of Internet bookshelf is proposed.

Section 2 and 3 give an overview of ontology and the semantic web respectively. The design of ontology based bookstore for m-commerce is presented in section 4. In section 5 conclusions are drawn.

2.0 Ontology

The term ontology has been in use for many years. Merriam Webster, for example, dates ontology circa 1721 and provides two definitions (i) a branch of metaphysics concerned with the nature and relations of being and (ii) a particular theory about the nature of being or the kinds of existence. Many authors propose slight variations in their definitions, in many cases to suit their research interests. [Grant, 2001] gives a widely accepted definition of ontology - a formal, explicit specification of a shared conceptualization. [Uschold and Gruninger, 2004] further explained each of the terms '*formal*', '*explicit*', '*specification*' and '*conceptualization*' as follows: A conceptualization refers to an abstract model of how people think about the world, usually restricted to a particular subject area or domain. Explicit means that the concepts and relations are given explicit names/definitions (terms and definitions). The *name* is a term and the definition is a specification of the meaning of the concept or relation. *Formal* means that specification is encoded in a language whose formal properties are well understood. Shared means that the main purpose of an ontology is generally to be used and reused across different applications and communities.

In summary, this definition can be expressed as a formal explicit description of concepts in a domain of discourse. This description is usually made up of axioms that define each thing, thus giving a model that represents the domain. This model can be used to reason about the objects in that domain and the relations between them. The use of ontology in representing knowledge and its relationship with epistemology is described in [Williams et al, 2004]

2.1 Current And Future Applications of Ontology :

Although ontologies have had a long history, they remained largely the topic of academic interest among philosophers, linguists, librarians, and knowledge representation researchers until somewhat recently. Currently Ontologies have been gaining interest and acceptance among computational audiences (in addition to philosophical audiences). [Guarino, 1998] provides a collection of fields that embrace ontologies, including knowledge engineering, knowledge representation, qualitative modeling, language engineering, database design, information retrieval and extraction, and knowledge management and organization. That collection put together in early 1998 did not include the web emphasis that is seen today.

In particular, ontology, as the shared and common understanding of the domain that can be communicated between people and application systems, has a significant impact on areas dealing with vast amount of distributed and heterogeneous computer-based information. These application areas include World Wide Web and Intranet information systems, complex industrial software application, knowledge management, electronic commerce and e-business. In this paper, we are concerned with marketing on the web with the use of mobile communication devices. There is increased need for shared semantics and a web of data and information derived from it. Scientist, researchers and regulatory authorities in genomics, proteomics, clinical drugs trials and epistemology all need a way to integrate these components. This is being achieved in large part through the adoption of ontologies [Shadbolt, 2006]. Generally, ontologies can be used to communicate between systems, people and organizations, interoperate between systems, and support the design and development of knowledge-based and general software systems [Williams et al, 2004]

3.0 Semantic Web and Semantic Web Services

3.1 Semantic Web

The Semantic Web aims to add a machine tractable, re-purposeable layer to compliment the existing web of natural language hypertext. In order to realize this vision, the creation of semantic annotation, the linking of web pages ontologies, and the creation, evolution and interrelation of ontologies must become automatic or semi-automatic processes. (i.e. in a distributed computation, semantic web aims at distributing data and services defined and linked in such a way that they can be used by machines, not just for display purposes, but for automation, integration and reuse of the data and services across various applications and heterogeneous users). [Berners–Lee et al, 2001] summarizes semantic web as a vision of a web of meaningful contents and services, which can be interpreted by computer programs. The emergence of the semantic web allows answers/results to be given to specified queries in the Internet; people can now add new facts/ideas to the Internet because they will be structured and maintained in such a way that machines can reason about them. A number of languages such as Extensible Markup Language (XML), Resource Description

Framework (RDF), RDF schema (RDF-S), and recently Web Ontology Language (OWL), etc., have been developed to facilitate the implementation of the concept of semantic web.

3.2 Semantic Web Services (SWS):

There are services on the web for users, defined through a service ontology, which enables machine interpretability of its domain knowledge. A web service is a software or program identified by a Uniform Resource Locator (URL), which can be accessed through the Internet through its exposed interface. The interface description declares the operations, which can be performed by the service, the types of messages being exchanged during the interaction with the service, and the physical location of ports, where information should be exchanged. For example, a web service for calculating the exchange rate between two money currencies can declare the operation getExchangeRate with two inputs of type string (for source and target currencies) and an Output of type float (for the resulting rate). A binding then defines the machine and ports where messages should be sent [Grant, 2001]

4.0 Ontology based Design of online bookstore for m–commerce:

We now use a domain, the selling and buying of books on the Internet to illustrate the application of ontology and SWS in m-commerce. The domain is an online bookstore. An online bookstore provides a means of allowing users to browse the web for a suitable offer to buy the books they want. The architecture of the proposed model is as presented in figure 1.

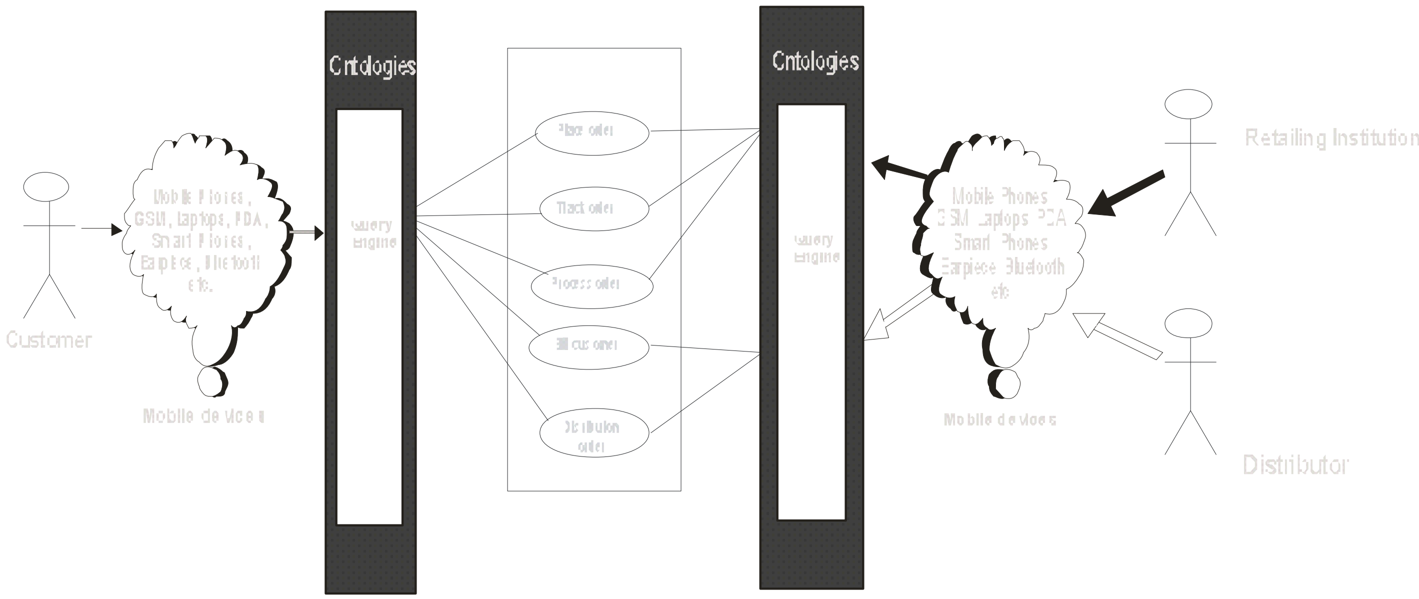


Figure 1: Conceptualized Model and Use Case Diagram

The architecture depicted in figure 1 shows two interface layers between the individual customer and the retail institution/distributor. The Mobile devices are used for communication and allow the user(s) to get services and run applications at any time and any place. The ontologies provide a means of describing the structure of semantics of objects/transactions in the domain and provides interoperability between the mobile devices of the customer and the retailing institution/Distributors. Between these layers are the possible transactions, which includes, placing order, tracking the order, processing the order, bill customer/receive bills and distribution order. The transactions are described with the use case diagram, which forms part of the design. The Use case diagrams describe what a system does from the standpoint of an external observer. The emphasis here is on what the system does rather than how it is done. The actors are the customer, the retail institution and the Distributing agent, who communicate through the specified Use Cases such as place order, track order etc., which depict some aspect of system functionality that is visible to an actor such as a customer. A semantic web service is provided to by ontologies actualizes these processes. The user adds information to the semantic web to accept the offer and gives details (such as name, shipping address, credit card information, etc.) of delivery via mobile devices. The buyer is notified of reception of his/her order leading to the completion of the transaction and the distribution and delivery of the requested book(s).

The type of ontology adopted in the design is the on2broker ontology described in [Fensel et al, 2001]. It has four main components, The query engine, and the information agent, inference engine, and database manager. The query engine receives query from the user and answers them by checking the content databases that were filled by the information agent and inference engines. The information agent is responsible for collecting factual knowledge from the web. Facts and ontologies are processed to derive additional factual knowledge by the inference engine. The backbone of the entire structure is the Database Manager, which receives and stores facts/knowledge from the information agent, exchanges the facts as input and output with the inference agent and provides facts to the query engine. The customer/consumer (the retail institution) interacts with the query engine, through mobile devices. The various activities possible on the internet bookshelf are presented in figure 2.

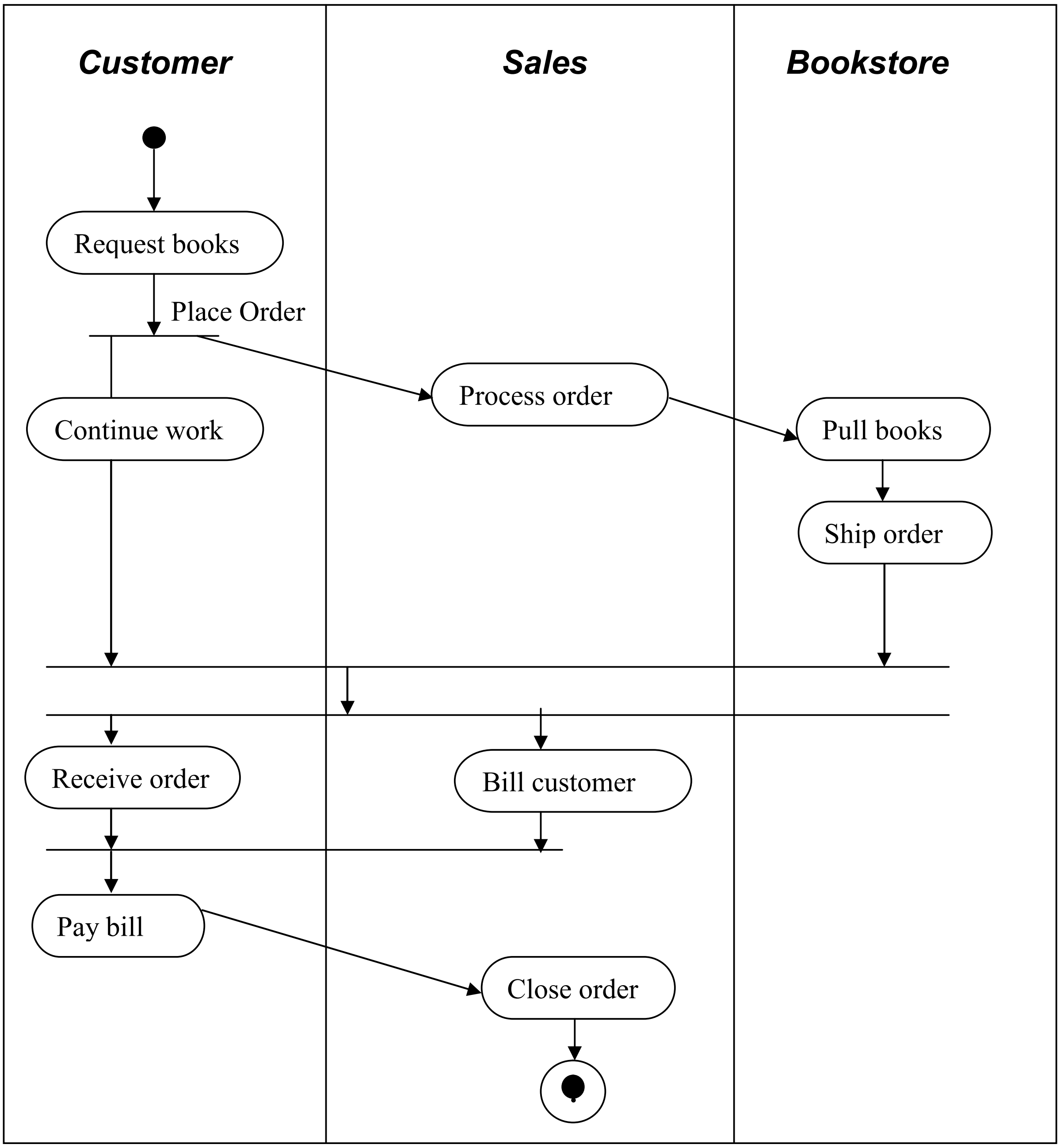


Figure 2: *Swimlanes for online bookstore design*

The Activity diagram for the online bookstore is shown in figure 2 while the class diagram is depicted in figure 3. The activity diagram represents a fancy flowchart with additional enhancements. It shows three objects; the customer, sales and bookstore, which interact with each other to accomplish the task of ordering a book from the bookstore. The activity diagram uses swim-lanes to show which objects take responsibility for which parts of the process. This can be represented by ontology showing objects and their relationships. The three classes involved; customer, sales and bookstore have activities they perform represented with rounded rectangles. The process begins at the black start circle at the top and ends at the concentric white/black stop circle at the bottom. The customer places an order by requesting books.

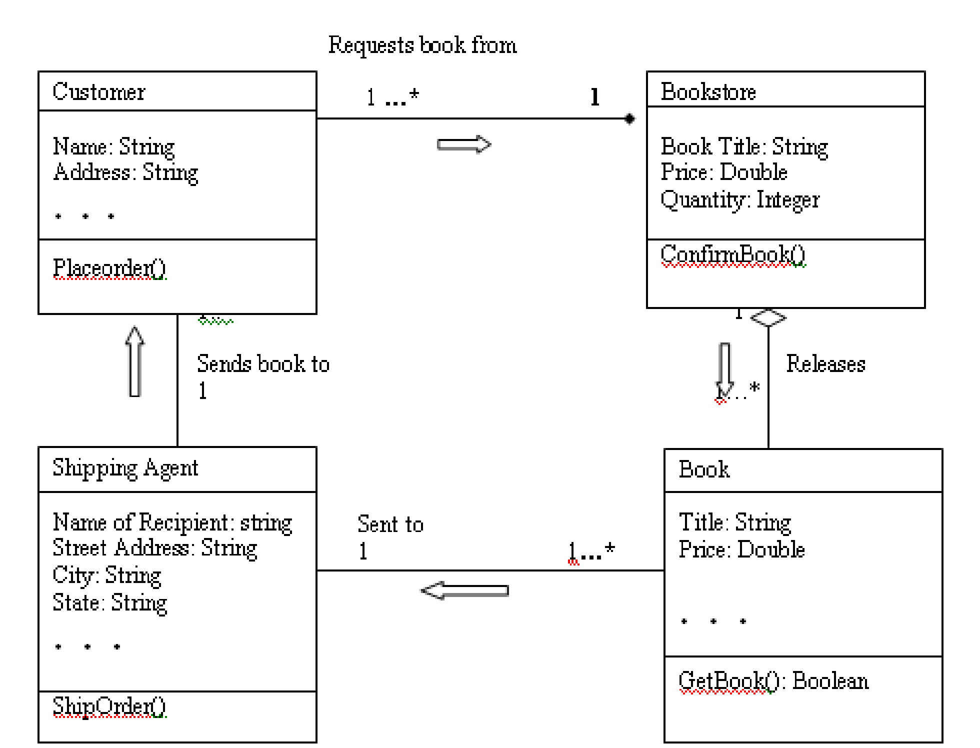


Figure 3: Class Diagram for Internet Bookshelf

The order is sent and processed in the sales and eventually on successful processing of the order; the requested books are pulled from the bookstore and shipped to the customer. The customer pays bill(s) to the sales and the transaction is closed after an order shipment and credit card billing is successful. As indicated in the figure 4.2, the customer can continue working while waiting for response; such as notification of reception of order placement , from the sales.

The class diagram depicted in figure 3, is used to describe the types of objects in a system and their relationships. Class diagrams give an overview of a system by showing the classes and the relationships among them, represented by ontology. It shows interaction between the customer and the bookstore as well as the shipping agent. Since a bookstore can be empty, there is an association relationship between the bookstore and a book. Many books can be found in one bookstore and no book may be found as well.

The book class has a GetBook() function of type Boolean that returns true if the requested book is found, otherwise returns false. A many to one relationship exists between the shipping agent and book class as well as the customer and shipping agent.
Each class has three parts:

i. a name e.g. Customer
ii. attributes e.g. Name, Address, Price etc.
iii. operations e.g. PlaceOrder(), GetBook() etc.

5. Conclusion

People engage in commerce by using or adopting the tools and technologies that are available at the moment. Also, Internet facilities and mobile devices have been used by business managers and customers and have created a good platform for them to communicate. Mobile devices are capable of getting services and runing applications at any time and at any where, especially while on the move. It also provides enhanced security (SIM card for mobile phones store confidential user information), convenience, accessibility and personalization. These factors have caused rapide acceptance of m-commerce leading to voluminous data, complex transactions and increased nature of relationship among data items and transactions; hence the need for effective tools to overcome these problems. This paper proposes a design based on ontologies and semantic web services for m-commerce applications in the domain of Internet bookshelf. This design enhances the security, accessibility, interoperability and communications between systems, devices, and objects and provides a framework for effective and easy m-commerce.

REFERENCES

Berners-Lee, T., Hendler, J., Lassila, O. (2001). The Semantic Web, Scientific American, Vol.284(4) pp 34-43.

Fensel, D., Angele Jurgen, Decker S., Erdmann M, Schnurr H., Staab S., S. Rudi and Witt A (2001): On2broker: Semantic-Based Access to Information Sources at the WWW. http://www.aifb.uni-karsruhe.de/www-broker.

Gary, P. S., (2003): Electronic Commerce, Canada: Kristen Duerr.

Grant, D., (2001). Internet Banking Nightmare: Couple. Sue after access to their fund was cut off 10 crucial days. Eastside J., Com (http://www.eastsidejournal.com)pp: 4-18.

Guarino, N. (1998): Some Ontological Principles for Designing Upper Level Lexical Resources. Proc. Of the First International Conference on Lexical Resources and Evolution, Granada, Spain.

Knospe, H. and Schwiderski-Grosche, S. (2002): Online Payments for Access to Heterogenous Mobile Networks. Ist Mobile & Wireless Telecommuncations Summit. Thessaloniki Greece.

Shadbolt, N. Hall, W. and Berners-Lee T., (2006): The Semantic Web Revisited. IEEE Intelligent Systems – www.computer.org/intelligent. pp 96-101

Tarasewich P., Nickerson R. and Warkentin M. (2002): Issues in Mobile e-Commerce. Communications of the Association for Information Systems. Vol. 8. pp 41-64.

Uschold, Michael and Gruninger, Michael (2004). Ontologies and Semantics for Seamless Connectivity. SIGMOD Record, Vol.33, No. 4.

Veijalainen J. Terziyan V. and Tirri H. (2003): Transaction Management for M-commerce at a Mobile Terminal. Proceedings of the 36th Hawaii International Conference on Systems Sciences.

Williams, E. E., Ayeni J. O., Z Lipsey (2004): Knowledge Representation in AI: From Epistemology to Ontology. The Journal of Computer Science and its Applications. Vol. 10 No.1 pp 127 – 140.

A MULTI-CHANNEL APPLICATION FRAMEWORK FOR CUSTOMER CARE SERVICE USING BEST-FIRST SEARCH TECHNIQUE

O.M. ADEYEYE, A.A. ATAYERO

Covenant University, Ota, Nigeria

micadeyeye@yahoo.com, atayero@ieee.org

Abstract

It has become imperative to find a solution to the dissatisfaction in response by mobile service providers when interacting with their customer care centres. Problems faced with Human to Human Interaction (H2H) between customer care centres and their customers include delayed response time, inconsistent solutions to questions or enquires and lack of dedicated access channels for interaction with customer care centres in some cases.

This paper presents a framework and development techniques for a multi-channel application providing Human to System (H2S) interaction for customer care centre of a mobile telecommunication provider. The proposed solution is called Interactive Customer Service Agent (ICSA). Based on single-authoring, it will provide three media of interaction with the customer care centre of a mobile telecommunication operator: voice, phone and web browsing. A mathematical search technique called Best-First Search to generate accurate results in a search environment.

Keywords: Artificial intelligence, Computer Languages, Internet, Multi-access communication

1.0 INTRODUCTION

With the present state of development in telecommunication in terms of bandwidth and speed, people now demand for various internet services via their mobile devices. Besides, they long for a more convenient way of using the web through voice interaction.

Owing to this global state of technological advancement, industries have ceaselessly continued to satisfy their customers through astounding applications. Typically, banks and mobile operators now provide different means of rendering services from human to human interaction (H2H) to human to system (H2S) interaction.

This paper has been segmented into five sections. Section 2 gives brief and concise information on WEB, Wireless Application Protocol (WAP) and VOICE architectures with an integrated diagram of the architectures. In addition, it presents information on multimodal and multi-channel applications, single authoring and multiple authoring, and the Best-first search technique. Section 3 examines design issues and practices. It describes how xml transformation is done and the N-tier models. In section 4, the framework is emphasized by indicating the features of each layer of the used N-tier model. Section 5 explains how the proposed framework will operate with the aid of a Use Case diagram and our conclusion is given in section 6.

2.0 BACKGROUND THEORY

2.1 WEB Architecture

The World Wide Web operates on Hypertext Transfer Protocol (HTTP). This is a client/server architecture whereby the server resides at one end and serves web pages to client at another end. A browser resides on the client and is used to interpret Hypertext Markup Language (HTML) codes passed by the server. Web pages or files are classified into static and dynamic. The static pages are fully interpreted at the client end while dynamic pages are executed at the server end and results passed to client. Such processing is carried out by an application server. While a web server acts as a container or storage for the web documents or files with extensions like .html, .asp, and .jpg, database server stores data like images in binary format, figures, and texts.

2.2 WAP Architecture

Internet access over mobile devices requires a packet switching network which is obtainable today with the emergence of technologies like General Packet Radio Service (GPRS), Enhanced Data for GSM Environment (EDGE) and many more. Where this network is in place, a WAP gateway is required to interface with the existing internet. The WAP gateway is also connected to a Base Transceiver Station (BTS) that provides wireless connection to the mobile devices referred to as clients. The lightweight protocol in use is called WAP [1]. While a browser such as Internet Explorer (IE) ™ or Netscape Navigator is required to interpret HTML codes passed to a Personal Computer (PC), microbrowser like Openwave or Nokia is required to interpret Wireless Markup Language (WML) codes passed to mobile phones.

2.3 VOICE Architecture

In voice architecture an application resides on a web server. Voice recognition is performed by an Automatic Speech Recognition (ASR) server and text to audio conversion is performed by Text-To-Speech (TTS) server. The VoiceXML interpreter executes the application according to the VoiceXML specification [2, 3, 16]. There are two common ways of implementing voice interaction namely directed dialogue and mixed initiative or Interactive Voice Response (IVR). The prevalent of the two is directed dialogue in which a user is replied by a system or application that gives a set of instructions in voice format. The application interacts with users based on Dual Tone Multiple Frequency (DTMF) pulses generated on pressing the phone key(s). In this case, speech synthesis and voice recognition are not involved unlike IVR which requires that the application captures the voice of the user and intelligently interacts based on keywords or words that match available grammars. Grammars are set of words which are matched with the spoken words supplied by the user. Where one or more words match, the application executes a set of predetermined instructions based on its algorithm. IVR applications are based on programming languages like Voice Extensible Mark-up Language (VXML), and Speech Application Language Tags (SALT).

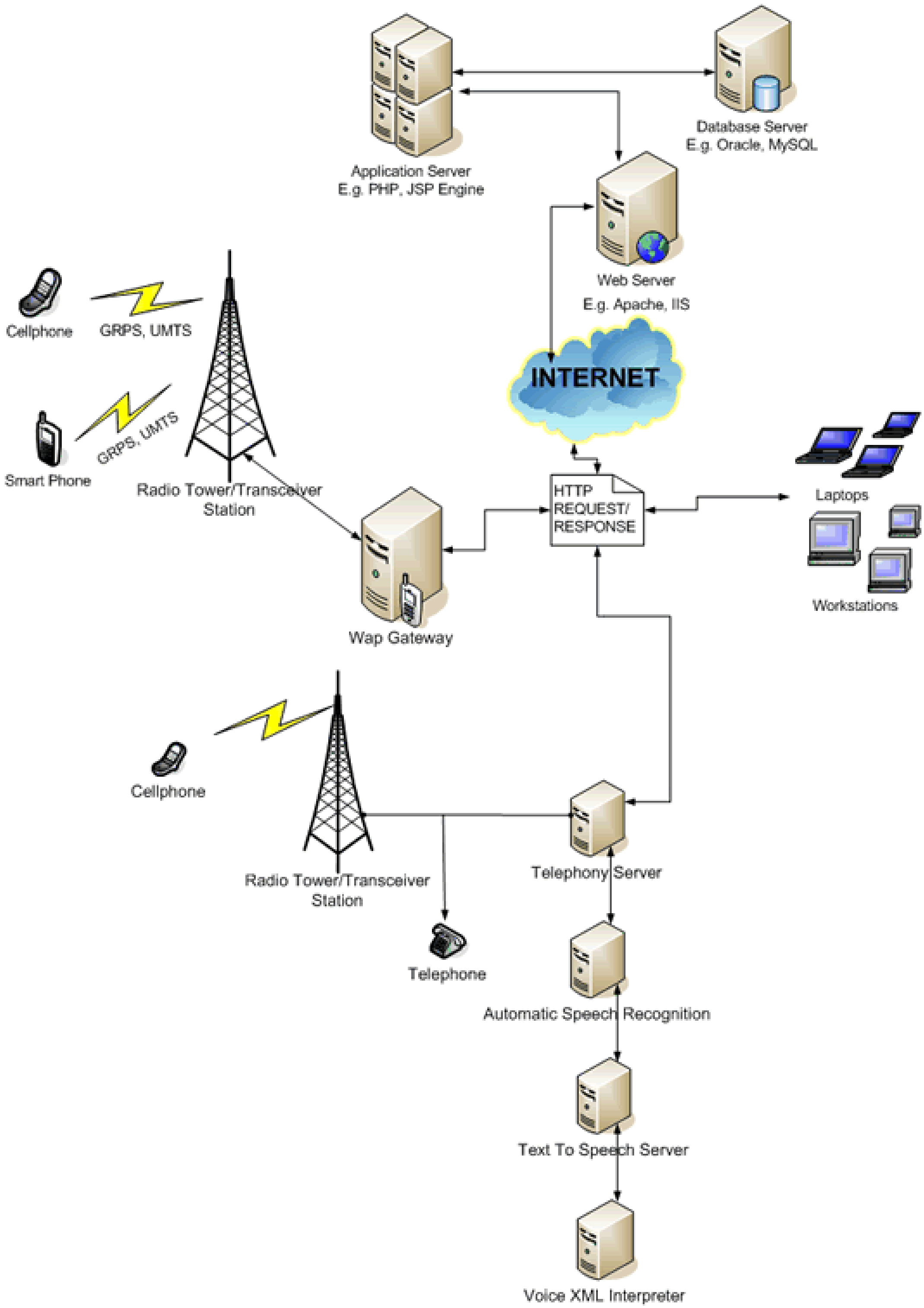


Figure 1 A WEB, WAP and VOICE integrated architecture

Fig.1 shows an integrated architecture that encompasses the WEB, WAP and VOICE networks. While PCs directly interpret HTTP request/response, mobile devices rely on the light-weight protocol known as WAP to present the information in compiled or binary format [1]. This is carried out through a gateway called WAP Gateway. For voice based interaction, the voice gateway may be made up of the telephony server, ASR server, TTS sever, and the VoiceXML Interpreter.

2.4 Multimodal Application vs. Multi-channel Application

Multi-channel access is the ability to access enterprise data and applications through multiple channels while multimodal access is the ability to combine multiple channels in the same interaction or session [3, 4]. X+V (XML+VXML) and SALT (Speech Application Language Tags) are the common tools for developing multimodal applications. While X+V focuses on combining web and telephony applications, SALT aims at enriching web applications with speech/telephony capabilities and turning them to multimodal. Also, SALT applications offer tightly-coupled multimodal browsing capability while X+V applications offer loosely coupled multimodal browsing in a more complete and structured manner.

Multimodal web applications are associated with many research challenges, mostly relating to multimodal interaction. One general usability requirement is to provide a highly improved experience to the end-user through robust, user-friendly applications [5, 6]. To this end, architectures capable of synchronizing and supporting voice and data simultaneously and seamlessly are required unlike Multi-channel where supporting data and voice are based on user's preference.

Network and computational requirements need to be taken into account in Multi-channel and Multimodal applications. An issue affecting the client devices is the presentation of different content formats enabling the various modalities, for both fixed and mobile devices. There are also issues relating to authoring and reusability of server-side components, application design, as well as the need for dealing with access and transformation of content at the system's back-end [5].

2.5 Multiple Authoring and Single Authoring

Multiple authoring provides simplicity because it relies on multiple representations of the target data for the supported modalities. Multiple-authoring means that each modality has its own presentation tier. From an application developer's point of view, in a multiple-authoring approach each modality is served from a separate server or web directory [5].

A single-authoring approach reuses common blocks of the different presentation tiers, toward a common presentation tier for all modalities. This is based on generating and handling multimodal content through appropriate processing of user requests, content acquisition and content transformation according to the various modalities.

Key benefits of multiple-authoring are that it facilitates content acquisition, minimizes content transformations and ease of extending existing content (e.g. HTML) to multimodal content. Scalability problems arise mainly because of redundant authoring which is alleviated by single-authoring. Single-authoring approach keeps content management tasks simpler and naturally tailors the application to multi-architecture [5].

2.6 The Search Technique

Artificial Intelligence (AI) is the study of how to make computers do things which, at the moment, people can do better [7]. Common AI techniques include question answering, generate and test, breadth-first search, deep-first search, best-first search and many more. We now consider the Best-First Search technique in detail.

2.6.1 Best-First Search

Best-first search is a way of combining the advantages of both depth-first and breadth-first search into a single method. Depth-first search is good because it allows a solution to be found without the need for expanding all competing branches, while Breadth-first search is good because it does not get trapped on dead-end paths [7].

2.6.2 Capabilities of Best-First Search Technique

By combining the merits of Depth-First Search and Breadth-First Search techniques, it has less memory usage (an advantage of Deep-First Search). Elaborately, it examines a node (domain of likely answers) only when an answer has not been found as shown in fig. 2. Because of this, the time taken to perform a search will be reduced. Agreeably, the processing speed of a query will be high. In addition, it will provide a solution if it exists. That is, it would not get trapped or terminated abruptly (an advantage of Breadth-First Search).

To implement this search technique, it will be required that existing answers are grouped into domains. Also, the system must intelligently narrow question asked to very few domains of possible answers.

2.6.3 Best-First Search Tree

1. Start with OPEN containing just the initial state.
2. Until a goal is found or there are no nodes left on OPEN do:

a. Pick the best node on OPEN
b. Generate its successors
c. For each successor do:

i. If it has not been generated before, evaluate it, add it to OPEN, and record its parent.

ii. If it has been generated before, change the parent if this new path is better than the previous one. In that case, update the cost of getting to this node and to any successors that this node may already have.

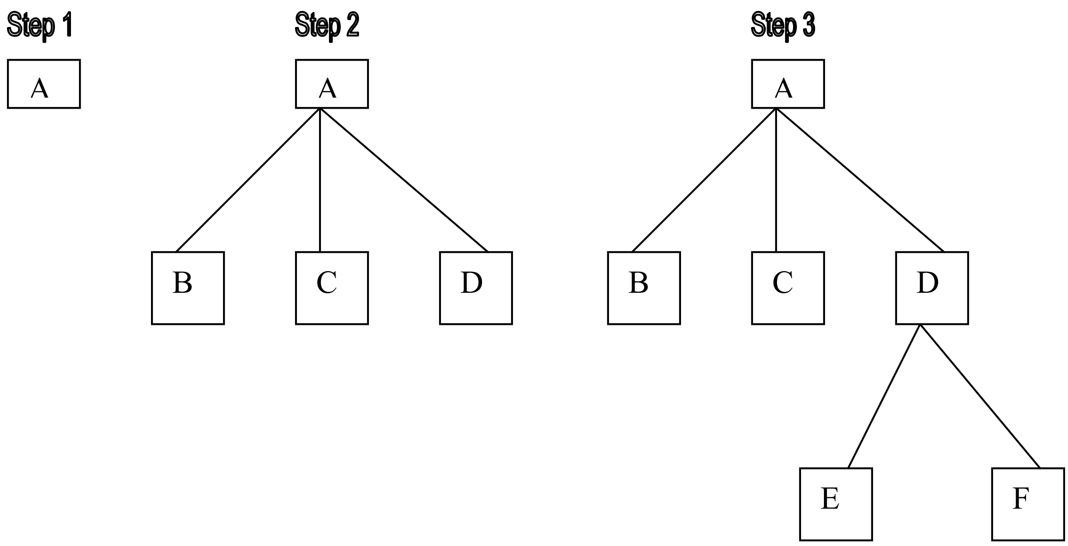


Figure 2 (Source: [7]) A Best-First Search Tree

3.0 DESIGN ISSUES AND PRACTICES

With a goal of developing a multi-channel application using single-authoring approach, basic design issues relating to n-tier application are considered. These design issues and practices include the XML/XSL Transformation and different layers in N-tier models. Their attendant benefits include ease of content management, and scalability to mention a few.

3.1 XML/XSL Transformation

Extensible Mark-up Language (XML) is a meta-markup language. It is a set of rules for creating semantic tags used to describe data. While HTML is used to specify the layout of a web page, XML is used to describe data [8, 9]. Extensible Stylesheet Language (XSL) is an xml-based language that can be used to manipulate, sort, and filter XML data. XSL language has been further split into three parts:

- Transformation (XSLT)
- Rendition (XSLF)
- XPATH

Extensible Stylesheet Language Transformations (XSLT) treats the document to be transformed as a set of nodes. An XSLT Stylesheet defines a set of rules or templates. When a template matches one of the nodes in the source document, it results in storing the output structure given by the template in a new document. XSLT uses the World Wide Web Consortium (W3C) XPATH specification language to query XML data. World Wide Web was created in October 1994 to develop common protocols that promote the evolution of the web and ensure its interoperability [10]. XPATH is strongly analogous to SQL and lets one specify complex rules to match nodes in a document [11]. A multi-channel application presents its content to the end users based on their connecting devices [5, 12]. XSL is an ideal tool for separating content and presentation. It can be used to overcome the inherent difficulties associated with the presentation of data to devices with limited capabilities as shown in fig. 3.

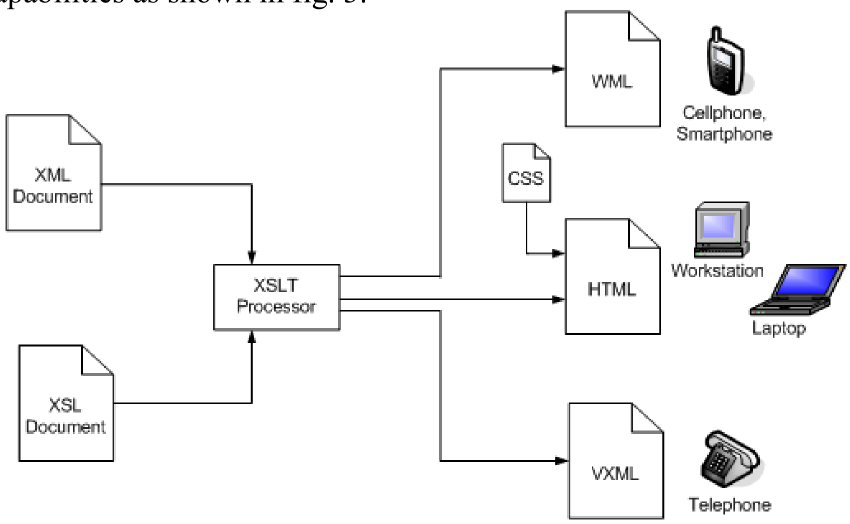
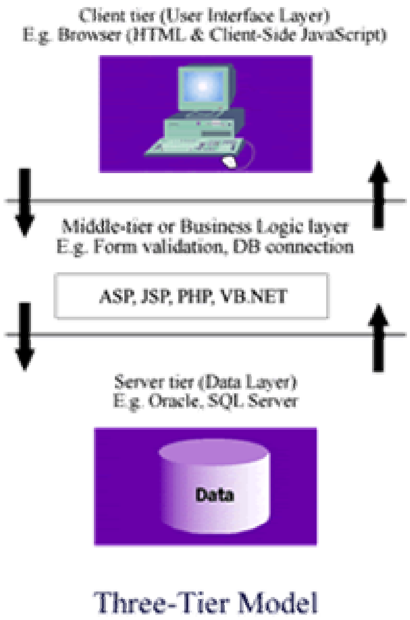


Figure 3 XML/XSL Transformation to HTML, WML and VXML

3.2 N-tier models

N-tier refers to the number of layers, into which a whole package or application can be classified. Better still, it can be referred to as Client-Server model where N>=2 [13]. Web application would perform better if developed with a five-tier model in mine rather than three-tier model. Five-tier model is an expansion on three-tier model as shown in fig. 4

Key benefits of N-tier applications include scalability to cope with future traffic and performance demand; and design of well structured, flexible, and vendor-neutral applications that are very easy to maintain [13].



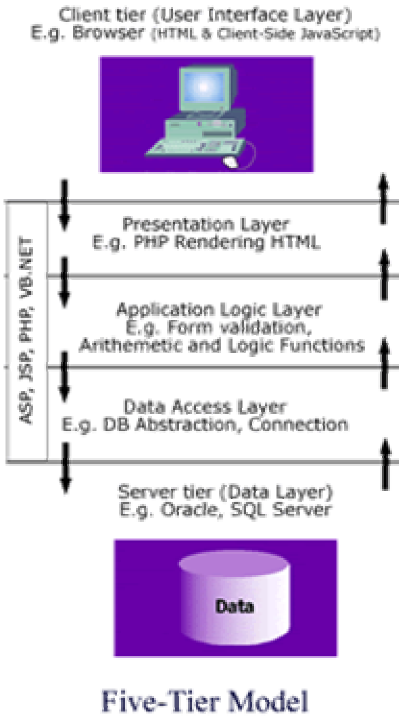


Figure 4 (Source: [13]) Three-tier and Five-tier models

4.0 THE PROPOSED FRAMEWORK FOR IMPLEMENTING ICSA

A framework for a multi-channel application called ICSA is introduced here. It employs a five-tier model as shown in fig. 5.

4.1 Data-Layer

It represents the database and web servers. Database server contains keywords to be looked for when question(s) are asked, appropriate or similar questions and possible answers to be presented.

4.2 Data Access Layer

This contains a server-side include file or script that connects to the database. It establishes a connection to the database and provides a connection identifier for query purposes. An Application Program Interface (API) that converts the entire database into xml files can also operate at this layer.

4.3 Application Layer

It is responsible for the arithmetic and logical operations that are performed with the application. It includes searching for keywords and storing new set of questions when answers can not be provided. An efficient search technique called Best-First Search is used to query the database.

Where XML files represent the database, XPATH (similar to SQL) can be used to locate answers within the XML files. Also, a third party API can be used to perform the search in XML files. A server-side script generates an xml file on-the-fly where the query results are stored.

4.4 Presentation Layer

This layer contains extensible Stylesheet files that can format the on-the-fly document into WML, HTML, and VXML files depending on the device or browser used to access the application. The appropriate format is determined by examining the HTTP request made by the browser. Alternatively, a third party program like HAWXY can act as a proxy which examines the user's connecting device or browser and presents appropriate message to it.

4.5 User Interface Layer

It features different browsers or connecting devices. Where telephone or mobile phone is used, the application converts its message to voice with the aid of TTS server. In case of phone browsing, the browser feeds on WML files while in web browsing, the PC feeds on HTML files additionally formatted by Cascading StyleSheet (CSS).

The flow process is also shown in fig. 5. When a user connects to the application either by dialing a number or typing in a web address, the index file on the webserver is passed. As an xml file, information in the header request determines what XSL file should be used to present the response as HTML form, WML form or voice dialogue. On submission, a script queries the database and generates an on-the-fly xml document. The xml document is styled by XSL, and if need be by CSS also before presenting the response to user.

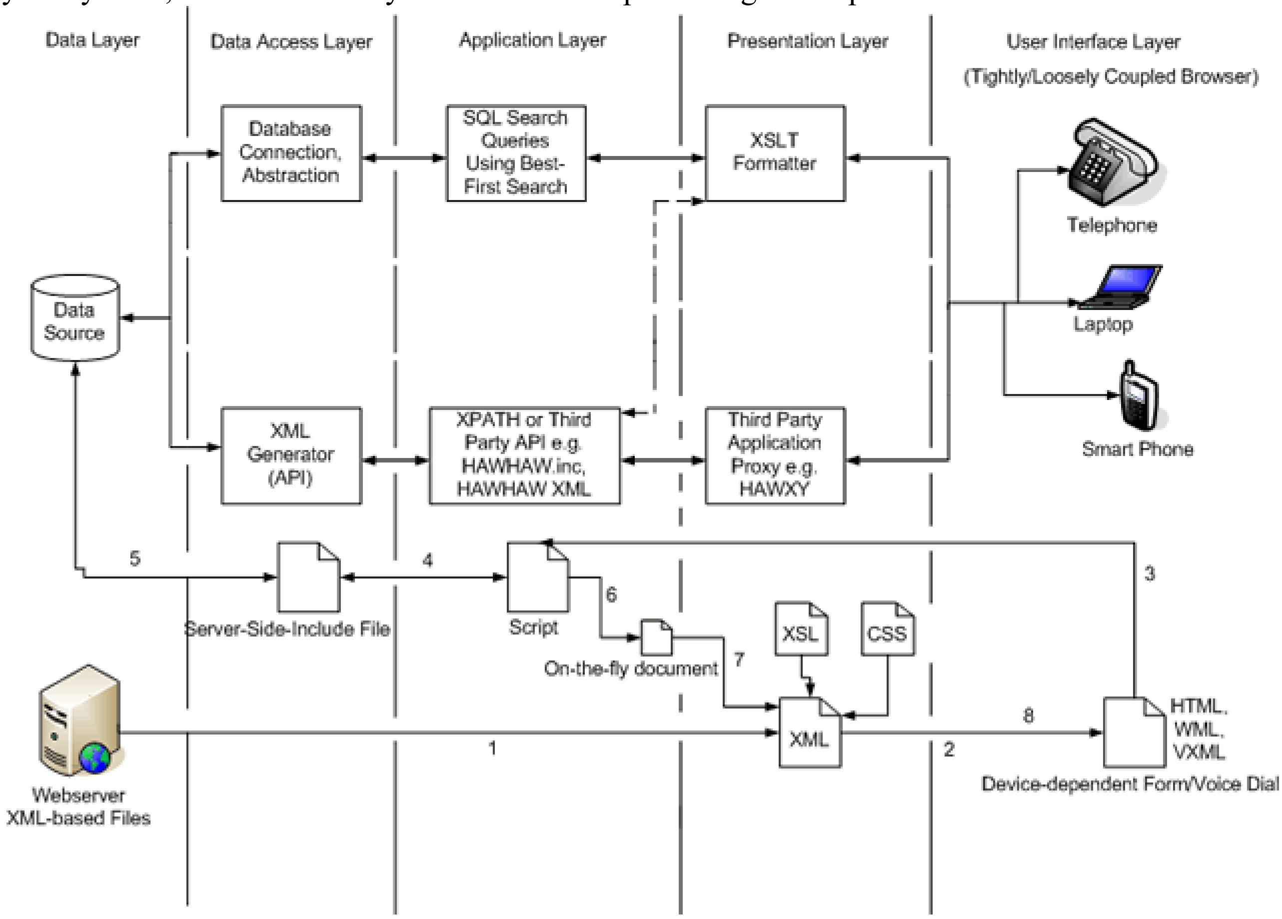


Figure 5 The proposed ICSA framework

This framework will require few XML files; separate XSL files for WEB, WAP, and VOICE presentations; and server-side scripting or application files to implement. It will provide a robust web application with ease of content management. In addition, it becomes easy to manipulate the xml files with the use of XSL. The framework gives room for scalability in order to add new features or extend it functionalities.

5.0 DISCUSSION

The use case diagram shown in fig. 6 pinpoints five steps involved in answering queries by the ICSA. A customer types in a question or a short sentence indicating his problem. On submission, the agent filters out the keywords, queries the database, and presents similar questions for the customer to pick the most similar to his. This first level of interaction helps in isolating the problem by searching for keyword(s) over a wide area (that is matching keywords in the entire database). This is referred to as horizontal domain sorting [14]. A confirmatory message is expected from the customer either by clicking a link (that is, the most similar question) or a voice response [15]. Thereafter, a thorough search called vertical domain sorting is carried out. This is an extensive knowledge on a narrow subject having isolated the problem from others. The most appropriate answer is presented. Should no answer exist, the question is stored for human response and a call is forwarded to a man or a message popup on a system at the customer service centre [2].

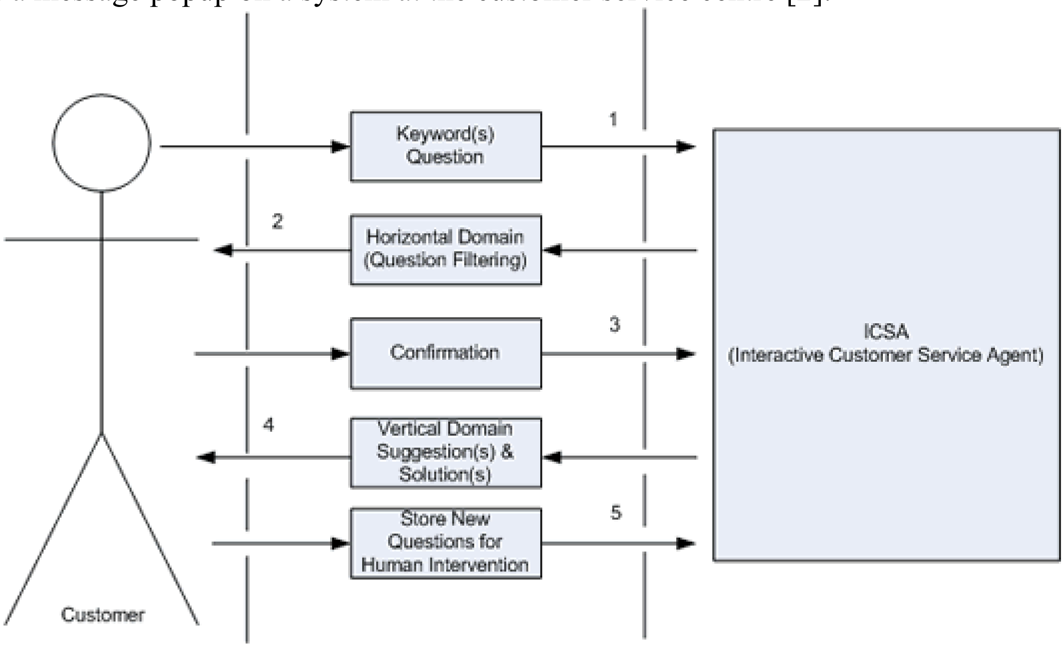


Figure 6 Use Case Diagram of the application

6.0 CONCLUSION

The proposed framework will produce a five-tier web application which is vendor-neutral, platform independent, scalable enough for future expansion, and robust. More so, it is a single-authoring programming paradigm using XML and can be implemented by coding in any server-side web language that supports XSLT. It can operate as a DTMF-based application over an analogue telephone network or IVR-based application over an IP-based or digital telephone network needless to say it provides three dedicated channels for customer interaction: phone, web and voice browsing.

Inherently, the application becomes reusable in future technologies or generations of fixed and mobile telecommunication. It will improve Customer Relationship Management (CRM) in telecommunication industry and any other firm that require 24/7/365 support for their invaluable services. It is germane to conclude that this proposition can boost productivity and sales for any business to consumer (B2C) model industries that implement this framework.

REFERENCES

[1] Chris Tull (2002), WAP 2.0 Development, Que Publication, USA, 1st Edition, pp 11-42.

[2] Atayero, A. A., Olugbara, O. O., Ayo, C. K., Ikhu-Omoregbe, N. A., "Design, Development and Deployment of an Automated Speech-controlled Customer Care Service System," *Proc. of the GSPx 2004 – "The International Embedded Solutions Event,"* Santa Clara, CA, September 27-30, 2004

[3] Vasiliki Demesticha, Jaroslav Gergic, Jan Kleindienst, Marion Mast, Lazaros Polymenakos, Henrik Schulz, L. Seredi: "Aspects of Design and Implementation of a Multi-channel and Multimodal Information System," *Proc. of the IEEE Conference on Software Maintenance*, November, 6-10, 2001, Florence, Italy (ICSM 2001), pp. 312-319.

[4] Scott A. (2000), "The rise of the Multimodal Web," *Proc. of the W3C/WAP Workshop: The Multimodal Web*, Hong Kong, 2000.

[5] Lazarous C.P. and John K. S.(2005), "Multimodal web applications: Design issues and an implementation framework", International Journal for Web Engineering and Technology, Vol. 2, No. 1.

[6] S. Ernest Chang: "The Design of a Secure and Pervasive Multimodal Web System," *Proc. Of the 19th IEEE Conference on Advanced Information Networking and Applications*, 2005.

[7] Elaine R. and Kevin K. (1991), Artificial Intelligence, McGrawHill, Singapore, International Edition, pp3-98.

[8] Ellen Pearlman and Eileen Mullin (2004), Programming the web using XML, New York, McGraw-Hill, 1st Edition, pp 1-9.

[9] Richard Light (1997), Presenting XML, USA, Sams.net, 1st Edition, pp 2, 202-209.

[10] Max Froumentin (2005), 'The World Wide Web Consortium,' Speech Technology, March/April 2005.

[11] Yannis Despotopoulos, John Soldatos, George Patikis, Lazarous Polymenakos, Jan Kleindienst and Jaroslav Gergie: "Accessing and Transforming Dynamic Content based on XML: Alternative Techniques and a Practical Implementation", *Proc. of the third international conference on Information Integration and Web based Applications and Services,* Linz Austria, 10-12 September 2001.

[12] M. Satyanarayanan, "Integrated Pervasive Computing Environments," *IEEE Pervasive Computing*, 1(2), 2002, pp. 2-3.

[13] Adeyeye Michael (2006), "Design and Implementation of a distributed Web-based application," Journal of Computer Science and its Applications, Vol. 12, No. 1.

[14] Boris Galitsky and Rajesh Pampapathi (2005), "Can many agents answer questions better than one?," First Monday, Vol. 10, No. 1.

[15] OnOn Hong (2004), VoiceXML Application Development Life Cycle: Building Usable Voice User Interface, http://devresource.hp.com/drc/resources/vxusability/vx_devlifecycle.pdf,

(April, 2005)

[16] William A. Yarberry, Jr.(2003), Computer Telephony Integration, Averbach Publication, USA, 2nd Edition, pp 1-107

RELIABILITY ISSUES OF GSM SECURITY PROTOCOL

O.S. OGUNDELE, O.S. ADEWALE AND B.K. ALESE
Department of Computer Science,Federal University of Technology,
P.M.B 704, Akure, Nigeria
email: oloruntoba78@yahoo.com

Abstract
Global System for Mobile Communication (GSM) began in 1982. The development of GSM marks the official beginning of second-generation cell phone technology. It provides several security mechanisms. These include anonymity, authentication, and confidentiality. The study therefore discussed the reliabilities issues of GSM security protocol which include weak crypto, SIM issues, the fake base station attack, and a total lack of replay protection and its effects on the future architecture of GSM technology.

Keywords: GSM, Security, Protocol

1.0 BACKGROUND TO THE STUDY

Prior to the early 1980s cell phones were expensive, completely insecure, and very large. These *first-generation* cell phones were analog and there were few standards and little or no thought was given to security. The biggest security issue with early cell phones was their susceptibility to *cloning* as stated in [1]. These cell phones would broadcast their identity in the clear when a phone call was placed, and this identity was used to determine who to bill for the phone call. Since this ID was broadcast over a wireless media, it could easily be captured and later used to make a "clone," or copy, of the phone. This allowed fraudsters to make free phone calls, which resulted in cheating on the legitimate consumers or the cellular phone companies. Cell phone cloning became a big business, with fake base stations created simply to gather IDs [2].

Global System for Mobile Communication (GSM) began in 1982 as Groupe Spéciale Mobile, but in 1986—as a tribute to the universality of three-letter acronyms—it was formally rechristened as Global System for Mobile Communications. The development of GSM marks the official beginning of second-generation cell phone technology [3].

1.1 GSM ARCHITECTURE

The general architecture of the GSM system is presented in figure 1 [1]:

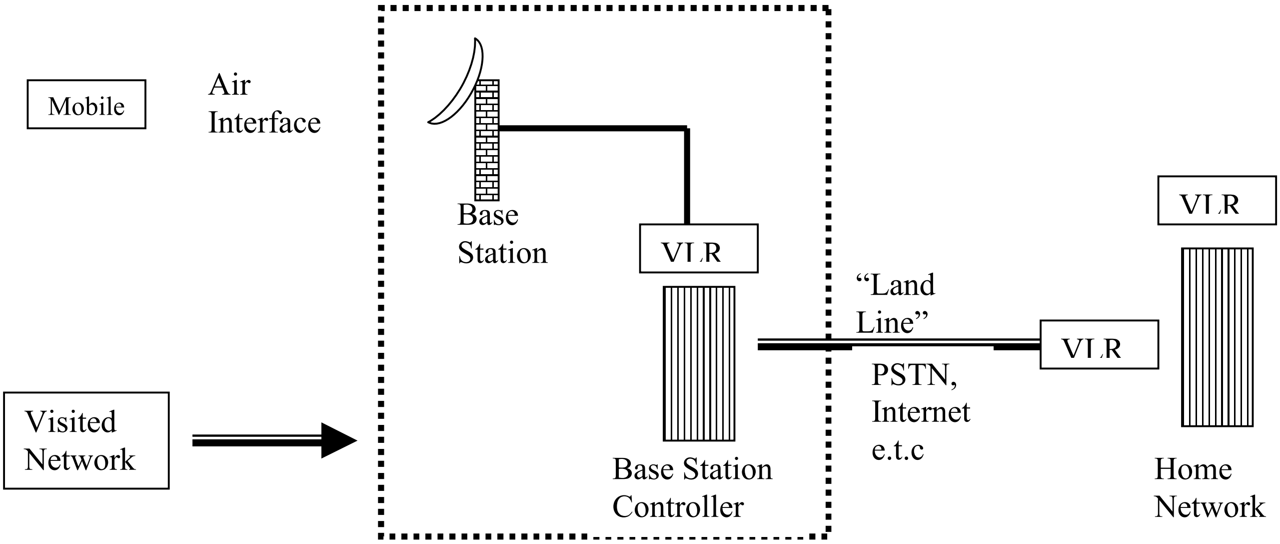


Figure 1: GSM Architectural Overview.

• The mobile is the cell phone.
• The air interface is the wireless transmission from the cell phone to a base station.
• The visited network includes multiple base stations and a base station controller, which acts as a hub for connecting the base stations under its control to the rest of the GSM network. The base station controller includes a visitor location registry (VLR) that is used to keep tabs on all mobiles currently accessing the GSM network via any of the base station controller's network of base stations.
• The public switched telephone network (PSTN) is the ordinary (non-cellular) telephone system. The PSTN is also referred to as "land lines" to distinguish it from the wireless network.
• The home network is the network where the mobile is registered. Each mobile is associated with a unique home network. The home network includes a home location registry, or HLR, which keeps track of the most recent location of all mobiles that belong to the particular home network. The authentication center (AuC) maintains the crucial billing information for all mobiles for which this particular home network is home.

Each GSM mobile phone contains a Subscriber Identity Module (SIM), which is a smartcard. On the SIM smartcard, there is an International Mobile Subscriber ID, or IMSI, which is used to identify the mobile. The SIM also contains a 128-bit key that is only known to the mobile and its home network. This key is universally known as Ki, as the standard notation. The purpose of the SIM smartcard is to provide an inexpensive form of tamper-resistant hardware. The SIM card also implements two-factor authentication, relying on "something you have" (the SIM itself) and "something you know" in the form of a four-digit PIN. However, the PIN is usually treated as an annoyance by most users, and it is generally not used.

Again, the visited network is the network where the mobile is currently located. A base station is one "cell" in the cellular system, whereas the base station controller manages a collection of cells. The VLR has information on all mobiles currently visiting the base station controller's territory. The home network stores a given mobile's crucial information, that is, its IMSI and Ki. The Home Location Register (HLR) keeps track of the most recent location of mobiles that have this home network as their home, while the AuC contains the IMSI-to-Ki mapping.

1.2 GSM SECURITY ARCHITECTURE

The primary security goals set forth by the designers of GSM were as follows [1];

- Make GSM as secure as ordinary telephones (the PSTN).
- Prevent cell phone cloning.

In particular, GSM was not designed to resist an active attack. At the time, active attacks were considered infeasible, since the necessary equipment was extremely costly. However, today the cost of such equipment is little more than that of a good laptop computer. The designers of GSM considered the biggest threats to be insecure billing, corruption, and similar low-tech attacks.

GSM attempts to deal with three security issues: anonymity, authentication and confidentiality [1]. The anonymity goal for GSM is to prevent intercepted traffic from being used to identify the caller. This particular feature is not especially important to the phone companies, but might be important to customers. Authentication, on the other hand is of paramount importance to the phone companies, since correct authentication is necessary for proper billing. The cloning problems with first generation phones can be viewed as an authentication failure. Confidentiality of calls over the air interface may be important to customers and probably to the phone companies too, at least from a marketing perspective.

1.2.1 ANONYMITY

GSM provides a very limited form of anonymity [1]. The IMSI is used to initially identify the caller. Then a Temporary Mobile Subscriber ID (TMSI) is assigned to the caller, and the TMSI is subsequently used to identify the caller.

In addition, the TMSI changes frequently. The net effect is that if an attacker captures the initial part of the call, where the IMSI is passed, then anonymity is not assured. But if the attacker misses the initial part of the call, then the anonymity is reasonably strong.

Although this is not a particularly strong form of anonymity, it may be sufficient for many practical situations where an attacker could have difficulty filtering the IMSIs out of a large volume of traffic.

1.2.2 AUTHENTICATION

Authentication is the most crucial aspect of GSM security. More precisely, authenticating the user to the base station is critical to the phone company; otherwise, they might not get paid for the service they provide. In GSM, the caller is authenticated to the base station, but the authentication is not mutual. The GSM designers apparently decided that it was not critical that the caller know he is talking to a legitimate base station. This was seen as a critical security oversight [1].

GSM authentication employs a challenge-response mechanism. The home network—which knows the Ki that corresponds to the caller's IMSI—is informed of the IMSI of the caller. The home network then generates a random challenge usually referred to as RAND (a random number between 0 and $2^{128} - 1$) and computes the "expected response," XRES =

A3 (RAND, Ki), where A3 is a hash function, and RAND and XRES are sent to the base station. The base station sends RAND, the challenge, to the mobile. The mobile's response is denoted SRES, where SRES is computed by the mobile as SRES = A3 (RAND, Ki). To complete the authentication, the base station verifies that SRES = XRES. One important feature of this authentication method is that the caller's key Ki never leaves its home network.

1.2.3 CONFIDENTIALITY

GSM uses a stream cipher to encrypt the data [1]. The reason for this choice is due to the high error rate, which is typically about 1 in 1,000 bits, in the cell phone environment. With a block cipher, each transmission error causes one or two entire plaintext blocks to be garbled (depending on the mode), while a stream cipher garbles only those plaintext bits corresponding to the specific cipher text bits that are in error. The GSM encryption key is universally denoted as Kc by convention.

When the home network receives the IMSI from the base station controller, the home network computes Kc = A8 (RAND, Ki), where A8 is a hash function. Then Kc is sent to the base station as part of the triple (RAND, XRES, Kc).

The base station receives the GSM triple (RAND, XRES, Kc) and completes the authentication as described above. If this succeeds, the mobile computes Kc = 8(RAND, Ki). The base station already knows Kc, so the mobile and base stations have established a shared symmetric key with which to encrypt the conversation. As mentioned above, the data is encrypted with a stream cipher. The keystream is generated as A5(Kc), where A5 is a stream cipher. As with authentication, the caller's master key Ki never leaves its home network.

2.1 GSM AUTHENTICATION PROTOCOL

The GSM protocol between the mobile and the base station is illustrated in Figure 2.

A few security concerns with this protocol are stated in [4] as follows:

- The RAND is hashed together with Ki to produce the encryption key Kc. Also, the value of RAND is hashed with Ki to generate SRES, which a passive attacker can see. As a result, it's necessary that SRES and Kc be uncorrelated. This is easy to achieve if the hash functions A3 and A8 are reasonably secure.
- It must not be possible to deduce Ki from known RAND and SRES pairs, since such pairs are available to a passive attacker. This is equivalent to a known plaintext attack on the hash function.
- It must not be possible to deduce Ki from chosen RAND and SRES pairs, which is equivalent to a chosen plaintext attack on the hash function. Although this attack might seem implausible, with possession of the SIM card, an attacker can choose the RAND values and observe the corresponding SRES values.

2.2 SHORTCOMINGS OF GSM SECURITY

There are cryptographic defaults in GSM and there are protocol flaws as well. The most devastating attacks on GSM are due to invalid security assumptions made by the original designers of GSM [1].

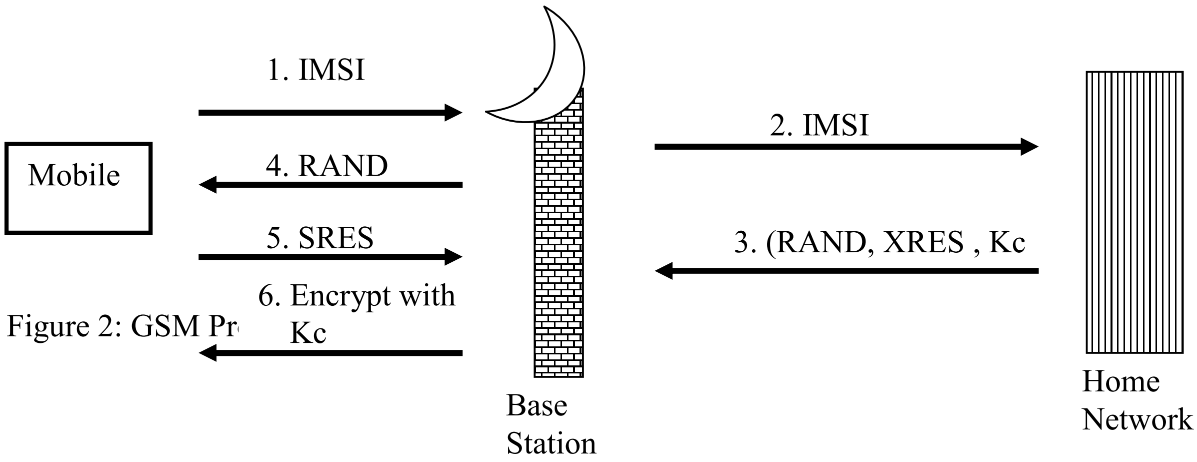


Figure 2: GSM Pr

2.2.1 CRYPTO FAULTS.

There are several cryptographic faults in GSM. The hashes A3 andA8 both rely on a hash function known as COMP128 [1]. The hash COMP128 was developed as a secret design, in violation of Kerckhoffs Principle. COMP128 was later found to be weak—it can be broken by 150,000 chosen plaintexts [5]. What this means in practice is that an attacker who has access to a SIM card can determine the key Ki in 2 to 10 hours, depending in the speed of the card. In particular, an unscrupulous seller of cell phones could determine Ki before selling a phone, and then create clones that would have their calls billed to the purchaser of the phone. There are actually two different forms of the encryption algorithm A5, which are known as A5/1 andA5/2. As with COMP128, both of these were developed in violation of Kerckhoffs Principle and both are weak. It has been shown in [6] and [7] that the A5/2 algorithm is weak and a feasible attack on A5/1 is known [8].

2.2.2 INVALID ASSUMPTIONS

There is a serious design oversight in the GSM protocol itself. A GSM phone call is encrypted between the mobile and the base station but not from the base station to the base station controller. Recall that the design goal of GSM was to develop a system as secure as the public switched telephone network (PSTN). As a result, if a GSM phone call is routed over the PSTN, then from that point on, no further special protection is required. In other words, GSM security is to protect the phone call over the air interface, between the mobile and the base station.

The designers of GSM assumed that once the call reached the base station, it would be routed over the PSTN to the base station controller. As a result of this assumption, the GSM security protocol does not protect the conversation when it is sent from the base station to the base station controller. However, many GSM systems today transmit calls between a base station and its base station controller over a microwave link [4]. Since microwave is a wireless media, it is possible for an attacker to eavesdrop on unprotected calls over this link, rendering the encryption over the air interface useless.

2.2.3 SIM ATTACKS

Several attacks were developed on various types of SIM (Subscriber Identity Modules) cards. One known attack called optical fault induction has been shown by [9] that it enables an attacker to force a SIM card to divulge its Ki by using an ordinary flashbulb. In another class of attacks, known as partitioning attacks, timing and power consumption analysis could be used to recover Ki using as few as eight adaptively chosen plaintexts [10]. As a result, an attacker who has possession of the SIM could recover Ki in seconds and, consequently, a misplaced cell phone could be cloned in seconds.

2.2.4 FAKE BASE STATION

Another serious fault with the GSM protocol is the threat posed by a fake base station. This attack exploits two flaws in the protocol. First, the authentication is not mutual. While the caller is authenticated to the base station (which is necessary for proper billing), the designers of GSM felt it was not worth the extra effort to authenticate the base station to the caller. Although they were aware of the possibility of a fake base station, apparently the protocol designers believed that the probability of such an attack was too remote to justify the (small) additional cost of mutual authentication. The second flaw that this attack exploits is that encryption over the air interface is not automatic. Instead, the base station determines whether the call is encrypted or not.

Some other major flaws with the GSM protocol is that it provides no replay protection. A base station (or a fake base station) can replay a compromised triple (RAND, XRES, Kc) forever. As a result, one compromised triple gives an attacker a key Kc that is valid forever. A clever fake base station operator could use a compromised triple to “protect” the conversation between the mobile and the fake base station so that nobody else could eavesdrop on the conversation. Denial of service is relatively easy in a wireless environment, since the signal can be jammed. GSM offers no special protection from such an attack also.

3.0 CONCLUSION

The two security goals set forth by the designers of GSM were to eliminate the cloning that had plagued first-generation systems and to make the air interface as secure as the PSTN. Although it is possible to clone GSM phones, it was never a significant problem in practice. However, it is perhaps more relevant to consider whether GSM achieved its security design goals.

The real problem with GSM security is that the initial design goals were too limited. The major insecurities in GSM include weak crypto, SIM issues, the fake base station attack, and a total lack of replay protection. In the PSTN, the primary insecurity is tapping, though there are others threats, such as attacks on cordless phones. In this light, GSM should probably be considered a modest security success.

It is worth mentioning that GSM is certainly a commercial success. Therefore it becomes imperative to raise some questions about the commercial significance of good security as depicted by GSM security protocol. Perhaps the GSM designers were more concerned with having a simple and “marketable” form of anonymity than providing a strong form of anonymity.

REFERENCES

[1]. Stamp Mark, "Information Security, Principles and Practice", Published by John Wiley & Sons, Inc., Hoboken, New Jersey. (2006),

[2]. Anderson R., "*Security Engineering"*, Published by John Wiley & Sons, Inc., Hoboken, New Jersey. 2001.

[3]. History of GSM, at http://www.cellular.co.za/gsmhistory.htm

[4]. Pagliusi P. S, "A contemporary foreword on GSM security", in G. Davida, Y. Frankel, and O. Rees, editors, *Infrastructure Security: International Conference—InfraSec 2002*, Bristol, UK, October 1–3, 2002, Lecture Notes in Computer Science 2437, pp. 129–144, Springer–Verlag, 2002.

[5]. GSM cloning, available at http://www.isaac.cs.berkeley.edu/isaac/gsm.htm

[6]. Barkan E., Biham E., and Keller N., "Instant cipher text-only cryptanalysis of GSM Encrypted communication", at http://cryptome.org/gsm-crackbbk.pdf

[7]. Petrovic S. and Fúster-Sabater A. (2000), "Cryptanalysis of the A5/2 algorithm", Available at http://eprint.iacr.org/2000/052/

[8]. Biryukov A, Shamir A, and Wagner D, "Real time cryptanalysis of A5/1 on a PC", Available at: http://home.in.tum.de/gerold/KryptDokumente/a5_Angriff/a51-bsw.htm

[9]. Skorobogatov S. and Anderson R., "Optical fault induction attacks", *IEEE Symposium on Security and Privacy*, 2002.

[10]. Rao J. R. et al. ,"Partitioning attacks: or how to rapidly clone some GSM Cards", *2002 IEEE Symposium on Security and Privacy*, May 12–15, 2002.

WIRELESS AND MOBILE NETWORKS

A.F. THOMPSON, B.K. ALESE AND O.S. ADEWALE

Department of Computer Science, Federal University of Technology, Akure
e-mail:ronkenira@yahoo.com

Abstract

Every new generation of technology challenges our world view and paradigms. For example, a paradigm shift occurred when people transitions from listening to the radio to watching television programs. Another example is when people went from using stand-alone personal computers to accessing the Internet. It is no surprise that mobility is causing yet another paradigm shift. Wireless technologies in the last decade have attracted unprecedented attention from wireless service providers, developers, vendors, and users. The breathtaking evolution of wireless technologies, services and business applications has resulted in a wide-scale deployment and usage of wireless and mobile networks. This paper focuses on the three main wireless and mobile networks, its requirements, basic elements associated with mobile and wireless networks, WLAN architectures, technologies and different types.

Keywords: Mobile, Wireless, Networks, LANs, WAN,

1.0 Introductions

Wireless and mobile networks are rapidly extending their capabilities. In addition to their increasing bandwidth and because of their flexibility and freedom they are becoming the communication infrastructure of choice. Wireless communication provides a user the capability of conducting commerce at anytime, with nearly anyone, from anywhere, using a mobile communication channel.

Wireless and mobile networks represent an increasingly important segment of networking research as a whole, driven by the rapid growth of portable computing, communication and embedded devices connected to the Internet. The currently deployed TCP-IP model of the Internet was originally designed for communication between wired PC's, mainframes and servers, and is clearly not optimized for wireless devices which may be expected to outnumber wired Internet terminals in the next ~5-10 years. Mobile and wireless service requirements have gradually been incorporated into networking standards (such as IPv6),

Wireless and mobile networks can be divided into three main categories:

- System interconnection.
- Wireless LANs.
- Wireless WANs.

1.0.1 System interconnection

System interconnection is all about interconnecting the components of a computer using short-range radio. Almost every computer has a monitor, keyboard, mouse, and printer connected to the main unit by cables. So many new users have a hard time plugging all the cables into the right little holes (even though they are usually color coded) that most computer vendors offer the option of sending a technician to the user's home to do it. Consequently, a short-range wireless network called Bluetooth was designed by a company to connect these components without wires. Bluetooth also allows digital cameras, headsets, scanners, and other devices to connect to a computer by merely being brought within range. No cables, no driver installation, just put them down, turn them on, and they work. For many people, this ease of operation is a big plus.

In the simplest form, system interconnection networks use the master-slave paradigm of Figure. 1a the system unit is normally the master, talking to the mouse, keyboard, etc., as slaves. The master tells the slaves what addresses to use, when they can broadcast, how long they can transmit, what frequencies they can use, and so on. (Andrew, 2003)

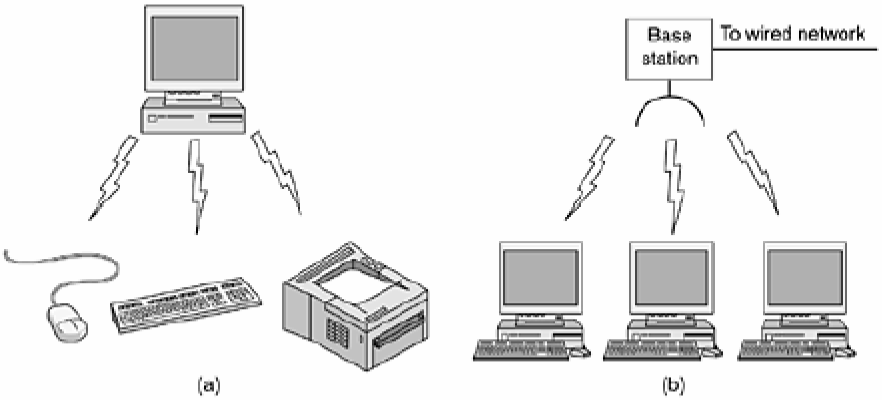


Figure 1. (a) Bluetooth Configuration. (b) Wireless LAN. (Andrew, 2003)

1.0.2 The wireless LANs.

These are systems in which every computer has a radio modem and antenna with which it can communicate with other systems. Often there is an antenna on the ceiling that the machines talk to, as shown in Figure. 2. However, if the systems are close enough, they can communicate directly with one another in a peer-to-peer configuration. Wireless LANs are becoming increasingly common in small offices and homes, where installing Ethernet is considered too much trouble, as well as in older office buildings, company cafeterias, conference rooms, and other places. There is a standard for wireless LANs, called IEEE 802.11, which most systems implement and which is becoming very widespread.

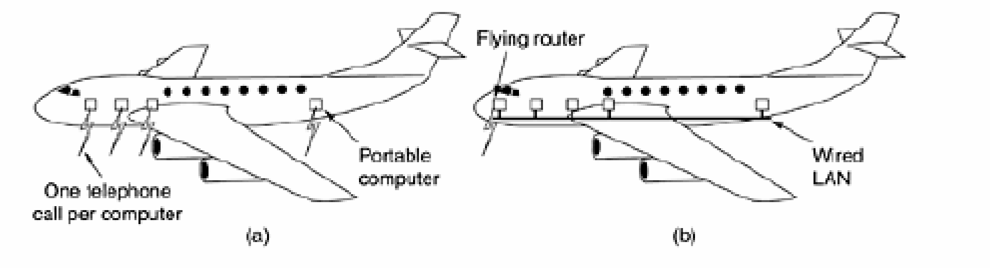


Figure 2. (a) Individual mobile computers. (b) A flying LAN. (Andrew, 2003)

1.0.3 The wireless WAN

The wireless WAN is Wireless Wide Area Network systems. The radio network used for cellular telephones is an example of a low-bandwidth wireless system. This system has already gone through three generations. The first generation was analog and for voice only. The second generation was digital and for voice only. The third generation is digital and is for both voice and data. In a certain sense, cellular wireless networks are like wireless LANs, except that the distances involved are much greater and the bit rates much lower.

Wireless LANs can operate at rates up to about 50 Mbps over distances of tens of meters. Cellular systems operate below 1 Mbps, but the distance between the base station and the computer or telephone is measured in kilometers rather than in meters.
In addition to these low-speed networks, high-bandwidth wide area wireless networks are also being developed. The initial focus is high-speed wireless Internet access from homes and businesses, bypassing the telephone system. This service is often called local multipoint distribution service
Almost all wireless networks hook up to the wired network at some point to provide access to files, databases, and the Internet. There are many ways these connections can be realized, depending on the circumstances. For example, in Figure 2(a), we depict an airplane with a number of people using modems and seat-back telephones to call the office. Each call is independent of the other ones. A much more efficient option, however, is the flying LAN of Figure 2(a). Here each seat comes equipped with an Ethernet connector into which passengers can plug their computers. A single router on the aircraft maintains a radio link with some router on the ground, changing routers as it flies along. This configuration is just a traditional LAN, except that its connection to the outside world happens to be a radio link instead of a hardwired line.

2.0 Wireless & Mobile Network Requirements: Wireless and mobile networks represent an active research and new technology development area. The rapid evolution of core radio technologies, wireless networks/protocols and application scenarios is summarized for reference in the technology roadmap given in Figure. 1 below. It can be seen from the chart that in addition to 2.5G/3G cellular data and WLAN systems developed during the 1990's, emerging wireless scenarios include personal-area networks, wireless peer-to-peer (P2P) , ad-hoc mesh networks, cognitive radio networks, sensor networks, RFID systems and pervasive computing .

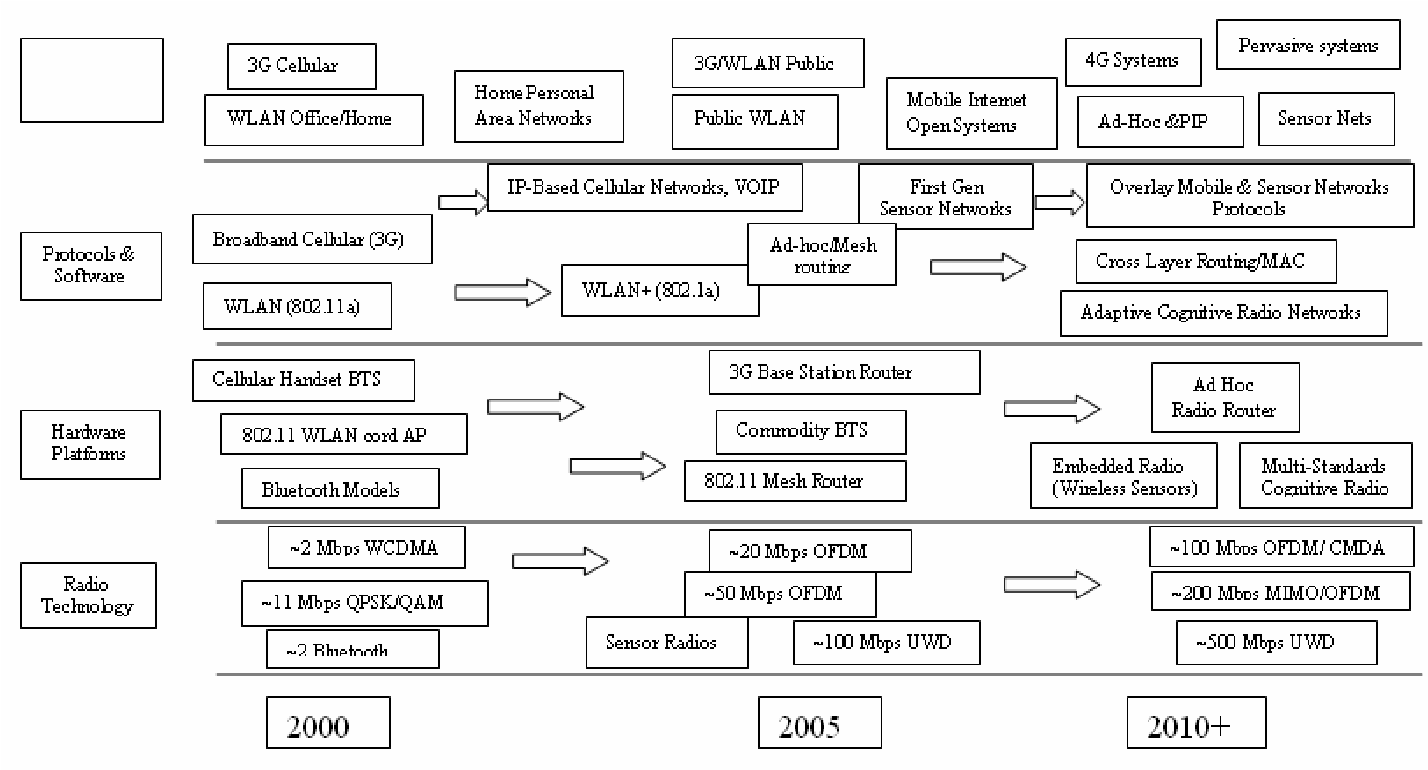


Figure 3: Wireless Technology Roadmap for the Period 2000-2010 (Per et al, 2002)

It is instructive to consider the specifics of several current and emerging wireless network scenarios to understand their networking requirements. Consider first the mobile data scenario where portable computers access the Internet on the move and therefore change their point of attachment to the network A third scenario that has received considerable attention in the research community (and more recently IETF and IEEE) is that of mobile ad-hoc (MANET) or mesh networks in which radio nodes organize themselves into multi-hop routing topologies that conserve power and take advantage of short-range radio links. The network must be capable of rapid topology discovery and self-organization, include support for mobile/wireless routers and preferably avoid loss of routing information across wired and wireless network segments. The ad-hoc network also requires improved security models as well as optional features such as geographic routing (Badri and Dragos,, 2002), epidemic dissemination and cross-layer protocol support. Significant research effort .Douglas et al., 2003 is currently focused on ideas for using cross-layer information from physical, MAC and network layers to improve the performance of routing and/or end-to-end transport protocols. In ad-hoc wireless networks, knowledge of physical layer parameters such as radio link speed and quality can provide important performance gains.

3.0 Basic Network Elements Associated with Wireless and Mobile Networks

This section provides a brief introduction to a few of the basic network components associated with the existing telecommunications infrastructure. It specifically discusses the existing mobile wireless network infrastructure components for TDM-based wireless networks, some of which eventually will be replaced by new IP-based components.

In the early 1980s, support for mobile wireless communications was introduced using cellular networks, which were based on analog technologies such as AMPS. Many of the telecommunications entities associated with cellular networks still play a vital role in today's wireless networks. As wireless communications technologies continue to progress and IP data networking is further integrated into the existing infrastructure, some of the functions of these entities might still exist within the network, but will be implemented in different and more effective ways.

The following network elements are part of a typical cellular telecommunications network:

- Public Switched Telephone Network (PSTN)

The PSTN is the foundation and remains the predominant infrastructure that currently supports the connection of millions of subscribers worldwide. The PSTN has several thousands of miles of transmission infrastructure, including fixed land lines, microwave, and satellite links. After the introduction of cellular telephone systems in the early and mid-1980s, and with the rapid development of mobile wireless communication services, the PSTN still provides the fixed network support using the Signaling System Number 7 (SS7) protocol to carry control and signaling messages in a packet-switched environment.

- Mobile Switching Center (MSC)

The MSC, usually located at the Mobile Telephone Switching Office (MTSO), is part of the mobile wireless network infrastructure that provides the following services:

- Switches voice traffic from the wireless network to the PSTN if the call is a mobile-to landline call, or it switches to another MSC within the wireless network if the call is a mobile-to-mobile call.
- Provides telephony switching services and controls calls between telephone and data systems.
- Provides the mobility functions for the network and serves as the hub for up to as many as 100 BSs.

More specifically, the MSC provides the following functions:

- Mobility management for the subscribers (to register subscribers, to authenticate and authorize the subscribers for services and access to the network, to maintain the information on the temporary location of the subscribers so they can receive and originate voice calls).

In GSM, some of the functionality of the MSC is distributed to the Base Station Controller (BSC). In TDMA, the BSC and the MSC are integrated.

- Call setup services (call routing based on the called number). These calls can be to another mobile subscriber through another MSC, or to a landline user through the PSTN.
- Connection control services, which determine how calls are routed and establishes trunks to carry the bearer traffic to another MSC or to the PSTN.
- Service logic functions, which route the call to the requested service for the subscriber, such as an 800 service, call forwarding, or voicemail.
- Transcoding functions, which decompress the voice traffic from the mobile device going to the PSTN and compresses the traffic going from the PSTN to the mobile device.
- Base Station (BS)

The BS is the component of the mobile wireless network access infrastructure that terminates the air interface over which the subscriber traffic is transmitted to and from a mobile station (MS).

In GSM-based networks, the BS is called a Base Transceiver Station (BTS).

- Radio Access Network (RAN)

The RAN identifies the portion of the wireless network that handles the radio frequencies (RF), Radio Resource Management (RRM), which involves signaling, and the data synchronization aspects of transmission over the air interface.

In GSM-based networks, the RAN typically consists of BTSs and Base Station Controllers (BSCs). User sessions are connected from a mobile station to a BTS, which connects to a BSC. The combined functions of the BTS and BSC are referred to as the Base Station Subsystem (BSS).

- Home Location Register (HLR)

The HLR is a database that contains information about subscribers to a mobile network that is maintained by a particular service provider. In addition, for subscribers of a roaming partner, the HLR might contain the service profiles of visiting subscribers.

The MSC uses the subscriber information supplied by the HLR to authenticate and register the subscriber. The HLR stores "permanent" subscriber information (rather than temporary subscriber data, which a VLR manages), including the service profile, location information, and activity status of the mobile user.

- Visitor Location Register (VLR)

The VLR is a database that is maintained by an MSC, to store temporary information about subscribers who roam into the coverage area of that MSC.

The VLR, which is usually part of an MSC, communicates with the HLR of the roaming subscriber to request data, and to maintain information about the subscriber's current location in the network.

- Authentication Center (AC)

The AC provides handset authentication and encryption services for a service provider. In most wireless networks today, the AC is collocated with the HLR, and is often implemented as part of the HLR complex.

4.0 WLAN Architecture

A WLAN architecture is built from stations and an access point (AP). The basic structure of a WLAN is the Basic Service Set (BSS). A BSS may either an *independent BSS* or an

infrastructure BSS.
In an independent BSS, the stations communicate with one another directly if they are within range of each other. These are sometimes referred to as *ad hoc* networks and generally last for a short time. These ad hoc WLANs are typically used for meetings and allow the participants to share data with one another. To participate in an ad hoc WLAN, the participants place the wireless network interface card (WNIC) of their devices into "ad hoc" mode. This mode allows a station to establish a connection with any other station in its proximity.
An infrastructure BSS requires the use of an access point (AP). The AP is used for all communications between stations. If one station wishes to send a transmission to another, the sending station sends its transmission to the AP. The AP then relays this transmission to the receiving station. This transmission requires two hops and will slow the WLAN. However the distance covered by the WLAN is increased by using the AP as a relay device. An important feature of the infrastructure BSS is the need for stations to associate to an AP. This feature can be used to set up a WLAN that has a form of restricted access. Figure 4 illustrates these concepts.

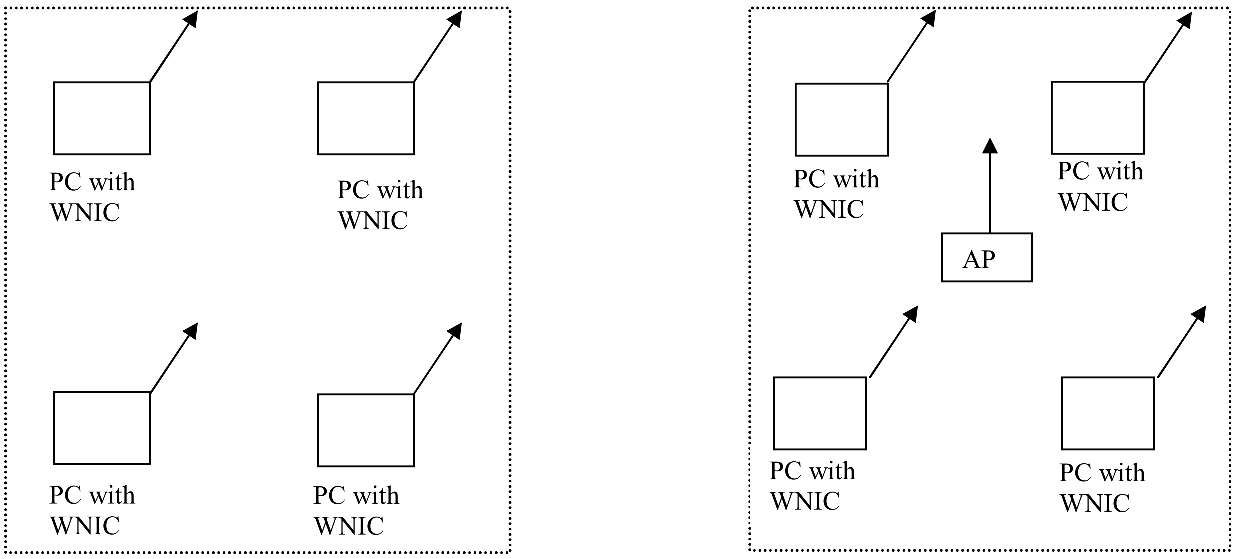


Figure 4: Basic Service Sets With and Without an AP (Robert , 2002)

These BSSs can be combined to form an *Extended Service Set* (ESS). An ESS is created by chaining together a number of BSSs by using a backbone network. The purpose of an ESS is to allow a station to have *transition mobility*. If a station has transition mobility, a user is able to roam from BSS to BSS and continue to be associated with a BSS and also have access to the backbone network with no loss of connectivity.

4.0.1 Wireless Physical Transport

The wireless signal that carries the data may be transmitted using electromagnetic waves in the either *radio frequency (RF) or infrared frequency (IR)* portion of the electromagnetic wave spectrum.
If RF Transport is used then the *Spread Spectrum* method is employed to generate the signal. The spread spectrum method expands the initial bandwidth and “spreads it out” in

order to use a portion of the expanded bandwidth for portion of the message. Two common variations of the spread spectrum technique are the *Frequency Hopping Spread Spectrum (FHHS) and the Direct Sequence Spread Spectrum (DSSS).*

When the FHSS variation of the spread spectrum is used, non-consecutive portions of the spread spectrum are used to transmit consecutive portions of the message. The transmitted message will be received out of order unless the receiver knows which portion of the spread frequency to tune to and how long to listen before hopping to the next frequency for a specific time period. An analogy would be listening to a song on the radio where the consecutive portions of the song are broadcast sequentially but on different stations. To hear the song correctly the listener would need to tune the stations in the correct sequence. The purpose of using FHSS is security and to reduce signal interference.

When the DSSS method is used, each portion of the message contains additional bits for error correction purposes - the message bits along with its redundant bits is called the “Chip Code” Because of the error correction bits, DSSS reduces the need to retransmit a signal and the result will be a more efficient use of the bandwidth.

If IR transport is used then the signal may be generated either as a *diffused signal* or a *point-to-point signal.*

A *diffused signal* can reflected off of existing surfaces such as a ceiling and that signal can be received by any device within range. A *point-to-point signal* is sent as beam to IR Switch that IR Switch relays the signals to next IR Switch and so forth.

RF is most commonly used of the two physical transport methods. In particular, the 802.11 standard employs the Industrial, Scientific, and Medical (ISM) RF band of the electromagnetic spectrum. This ISM band is specified as:

. • the I-Band from 902 MHz to 928 MHz,

. • the S-Band from 2.4GHz to 2.48GHz, and

. • the M-Band from 5.725GHz to 5.85GHz.

These bands are unregulated since they are used with low power. However, operating at low power limits the distance at which these signals can be detected. For example, depending on circumstances, using the S-band with a bandwidth of 1Mbps the distance varies anywhere from 300 feet indoors to 1500 feet outdoors.

5.0 Wirelesses and Mobile Technologies

The technologies related to wireless communication can be complex to differentiate. Wireless technology has been around for a while; however, there has been a relatively recent and rapid surge in the evolution of new wireless standards to support the convergence of voice, video and data communication. Much of this rapid evolution, or revolution, is a result of people seeking ubiquitous and immediate access to information and the assimilation of the internet into business practices and for personal use. People "on the go" want their internet access to move with them, so that their information is available at anytime, anywhere.

There are many factors that can be used to characterize wireless technologies:

- Spectrum, or the range of frequencies in which the network operates
- Transmission speeds supported
- Underlying transmission mechanism, such as frequency division multiple access (FDMA), time division multiple access (TDMA), or code division multiple access (CDMA)
- Architectural implementation, such as enterprise based (or in-building), fixed, or mobile

In addition, the mobile wireless technologies (such as Global System for Mobile Communications (GSM), TDMA, CDMA) are differentiated by a number of different factors, including some of the following:

- Control of the transmitted power
- Radio resource management and channel allocation
- Coding algorithms
- Network topology and frequency reuse
- Handoff mechanisms

6.0 Types of Wireless and Mobile Networks

Some wireless and mobile networks are discussed below.

6.0.1 Ad hoc networks

An ad hoc network is a collection of wireless mobile nodes dynamically forming a temporary network without the use of any existing network infrastructure or centralized administration. Conferencing, home networking, emergency services, personal area network and embedded computing applications are some of the potential applications for ad hoc networking.

Some open research issues in this field are quality of service, scalability, security and power control. The combination of this technology with the active network one, because they both focus on the flexibility of the network, its decentralized start up and maintenance, seems also challenging.

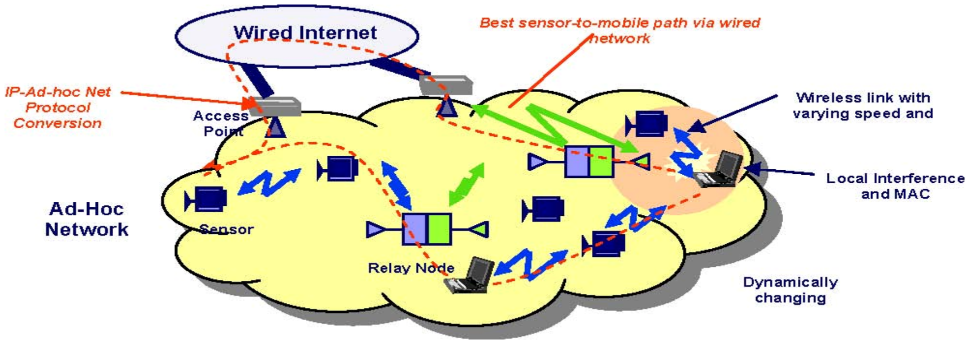


Figure 5: Ad-Hoc Mesh Network Scenario (MIT RoofNet Project, 2006)

6.0.2 Sensor Networks

Sensor networks, represent another major class of emerging wireless scenarios in which end-user devices are small embedded processors with processing and communication constraints due to limited battery power and size/cost. Such networks typically involve a multi-hop wireless ad-hoc access network similar to that shown in Figure 6 in order to deal with power limitations and the possible lack of wireless coverage in the area served. Most sensor nets also involve a hierarchy of network nodes ranging from tiny, low-power sensor devices (e.g. Berkeley Motes) (Wireless Sensor Networks), radio relay nodes (such as WINLAB's ad-hoc forwarding node (Sachin et al, 2005) and wired network gateways (such as the Intel Stargate node (Stargate)), as shown in Figure 6 below.

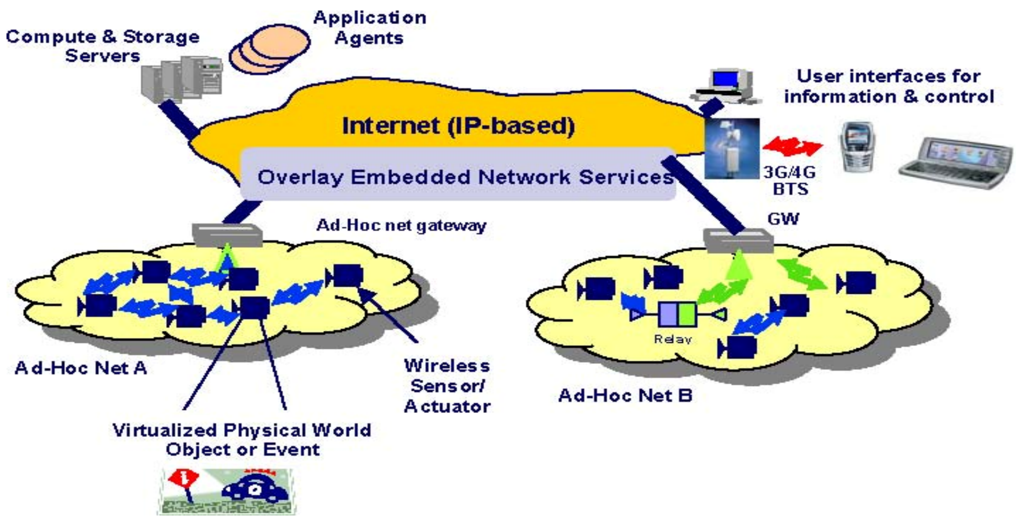


Figure 6. Embedded Wireless Network Scenario (MIT RoofNet Project, 2006)

In sensor networks, protocols at the low-tier nodes need to be optimized for power conservation and processing power limitations, so that conventional end-to-end Internet protocols such as TCP generally do not apply. Sensor systems with a hierarchy of network nodes as shown in Figure 6 can benefit from in-network processing and storage for data aggregation, search or caching that reduces processing and storage requirements at the sensor end points. This requirement is fundamentally different from earlier Internet design assumptions and needs to be factored into any future network architecture. In addition, sensor nets involve content- and location-aware applications in which end-user network addresses may not be known in advance so that the network needs to provide a dynamic binding service based on attributes or function. Content-based packet multicasting that delivers data to users based on interest profiles and queries is also a useful feature in such networks (Max et al, 2004). Overlay network services have been proposed as a means for providing such "data centric" services.

6.0.3 Personal-Area Networks

A Personal Area Network (PAN) is a network of devices that are closely associ-ated to one person. These kinds of devices may be attached to a person's clothing or carried around in a purse. In rather exotic visions of the future, these devices include virtual reality devices attached around the head and devices oriented toward the sense of touch. All the devices in a PAN may be attached to the Internet, but they will very likely have to communicate with each other. Because most devices in the same PAN will have an almost fixed position with respect to each other, mobility is not an important factor inside one PAN. However, when interactions between several PANs are needed, mobility can suddenly become much more important. To establish communication between moving PANs, ad hoc network technology can be used.

Personal area network is not only body but also protocol aspects of networking in the personal domain, digital home and other smart places like shopping centers, public and private transportation, ad-hoc communities, and networked (i.e., infrastructure provided) services (see Figure 7). In addition, the solutions should also work in a reasonable manner in non-smart places that do not have high-bandwidth connectivity or any connectivity at all.

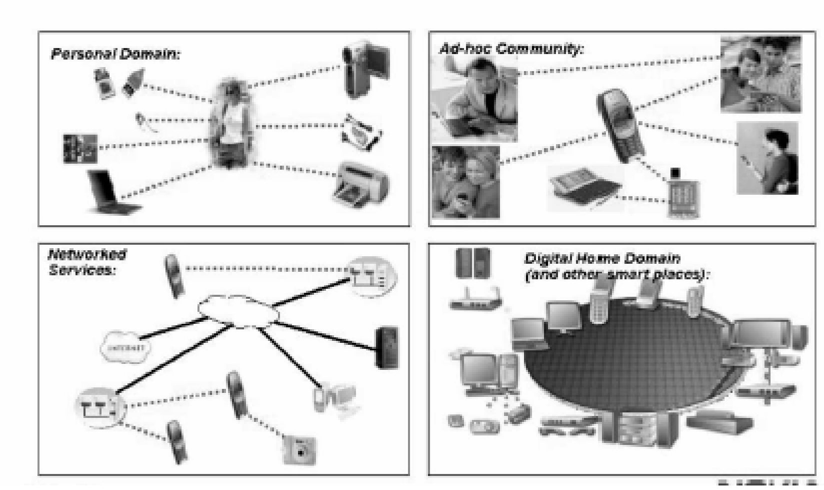


Figure 7: Personal Networking Domains (Andrew, 2003)

6.0.4 Pervasive Computing

Pervasive computing is characterized by a high degree of heterogeneity: devices and distributed components are from different vendors and sources. Support of mobility and distribution in such a context requires open distributed computing architectures and open protocols. Openness means that specifications of architectures and protocols are public documents developed by neutral organizations. Key specifications are required to handle mobility, service discovery, and distributed computing. Pervasive computing can be defined as availability of software applications and information anywhere and anytime (Birnbaum, 1997.) Personal Digital Assistants (PDAs) and cell phones are the first widely available and used pervasive computing devices. Devices and users are wireless and mobile. From an architectural point of view, applications are non-monolithic, but rather made of collaborating parts spread over the network nodes.

Context-aware components can sense who you are, where you are, and what you are doing and use that information to adapt their services to your needs. Mobility and services on demand are greatly impacted by the location of the devices and the requested services. Examples range from relatively rudimentary device following services such as phone call forwarding to the location of the device, to more complex issues of detecting locations of available services and selecting the optimal location for obtaining the services, such as printing services.

7.0 Conclusion

Computer networks can be used for numerous services, both for companies and for individuals.

Wireless networks are becoming extremely popular, especially wireless LANs. Networks provide services to their users. Mobile and wireless with various networks characteristics, technologies, speeds, and applications were highlighted in this paper. Future trends in this technology were brought to light.

8.0 References

Andrew S. Tannbaum, Computer Networks 4th Edition, 2003.

Badri Nath and Dragos Niculescu, "Routing on a curve", HotNets-I, Princeton, NJ, October 28-29, 2002.

Birnbaum J., Pervasive information systems, Communications of the ACM, 40(2): 40–41, Feb-ruary 1997.

Max Ott, Leslie French, Rajeev Mago, and Dhananjay Makwana. "XML-based semantic multicast routing: An overlay network architecture for future information services", In Proceedings of IEEE Global Telecommunication Conference (Globecom'04, Dallas, TX., USA, 2004.

MIT RoofNet Project, http://www.pdos.lcs.mit.edu

Per Johansson, Rohit Kapoor, Manthos Kazantzidis, Mario Gerla, "Personal Area Networks: Bluetooth or IEEE 802.11?", International Journal of Wireless Information Networks, Special Issue On Mobile Ad Hoc Networks (MANETs) Standards, Research Applications,

April 2002.

Robert J. Boncella , Wireless Security: An Overview, Communications of Association of Information Journal, volime 9, 2002, 269-282.

Sachin Ganu, Lalit Raju, Bhaskar Anepu, Suli Zhao, Ivan Seskar, and D. Raychaudhari. "Architecture and Prototyping of an 802.11-based Self-organizing Hierarchical Ad-hoc Wireless Network (SOHAN)" In Fifteenth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, volume 2, pp 880–884, 2004.

DEVELOPMENT AND SIMULATION OF AN ALGORITHM FOR DIMESIONING UMTS AIR INTERFACE FOR COVERAGE AND CAPACITY PREDICTION

O. Olabiyi, K. Ayodele, O. Ilori and K. Kehinde, MEMBERS, IEEE

Department of Electronic and Electrical Engineering
Obafemi Awolowo University, Ile Ife, Osun state, Nigeria.
e-mail: engr3os@gmail.com

Abstract

In the move from Global System of Mobile Communication (GSM) to Universal Mobile Telecommunication System (UMTS), the shift from Time Division Multiple Access (TDMA) technique to Wideband Code Division Multiple Access (WCDMA) resulted in network planning requirements being changed. The WCDMA system is interference limited, and this leads to a scenario where as more users are admitted into the network, the cell size shrinks. Therefore, coverage and capacity planning is done simultaneously. Estimating the system capacity requires both system and link level simulations, which requires significant computational resources. Hence, there is need for a fast and efficient algorithm that could be used for UMTS network deployment. This work gives a simple algorithm that can be used in estimating number of users that a Node B can serve considering both single and mixed service scenarios. The Hata-Okumura model was employed for coverage prediction. Uplink Load Equation was employed in determining the capacity and also address coverage-capacity trade offs with assumption that the network is uplink limited. The Load Equation was modified to account for the imperfect power control in the air interface. The effect of imperfect power control on system capacity was also simulated and discussed. Erlang model was then used to formulate call admission control and calculate the soft blocking probability. This algorithm, because of its simplicity can be made as an add-in in spreadsheet and can be used as dimensioning tool for planning UMTS network.

1 Introduction

A typical WCDMA system can either be coverage or capacity limited. This is due to the system been interference limited, in that as more users are allowed in the cell area, the noise level increases and therefore reduces the cell size. This work assumes that WCDMA radio interface is uplink limited with the mobile station transmit power being the major limitation, therefore only reverse link capacity is considered. Although recent trend has shown that WCDMA can be downlink limited this is due to the asynchronous nature of the radio interface with subscribers downloading at very high data rate and a lot of work has been done in this respect.

WCDMA uplink capacity, assuming non orthogonal uplink users and no multi-user detection is inversely related to the cell range, therefore there is always a compromise when planning both coverage and capacity. This compromise shown in [1] is depicted in the flowchart in Figure 1.

WCDMA planning starts from coverage determination by determining the maximum allowable pathloss. The pathloss can be used to derive the approximate cell range from existing propagation models. The link budget is used in determining the allowable pathloss and contains a term called noise rise. There exist a relationship between a given noise rise, load factor and the associated number of users which may be calculated using a simple uplink capacity technique called load equation as given by [2]. This is modified for imperfect power control for practical scenario resulting from Multipath propagation as described by [3].

The interference contribution of users in any cell, and the out–of –cell users within the network, result in a 'noise rise' relative to the thermal noise at the base station. This noise rise coupled with other link budget parameters, limits the maximum range. Since UMTS service is demand driven, the capacity is first estimated for a single service and this is then extended to other services. Depending on the kind of services, traffic models are applied in order to calculate the Erlang capacity and this is then used for final dimensioning.

The paper is divided into five sections. Section 2 discusses the propagation loss model and link budget. Section 3 describes the basic dependencies of radio interference load for uplink direction. Section 4 presents Erlang analytic models applied to blocking probability calculation for uplink direction with infinite source population. The following section discusses the result and the final concludes on the result.

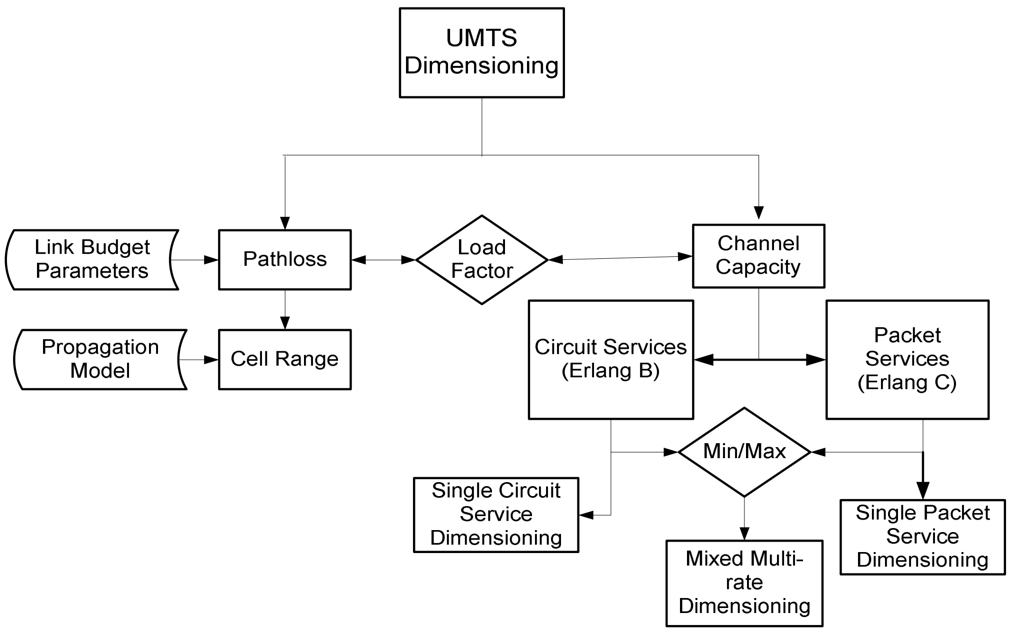


Figure1: The flowchart of coverage and capacity dimensioning of UMTS network

2 Coverage Prediction

2.1 Propagation loss model

Urban macro-cell environment: COST 231-Hata-Model

Hata-Okumura model [4] [5] is used to determine the range (coverage) of a cell covered by a base station. It is empirical in nature and hence much faster than other deterministic model.
Input parameters include frequency, *f in MHz*; distance between BS and MT, *d in km*; BS antenna height, *hBase in(m)*; and MT antenna height, *hMobile in (m)*. The following expression is valid for urban scenarios, but it can be extended for other scenarios by use of the respective correction factors of the Okumura et al. model:

$$L_b = 46.3 + 33.9\log(f) - 13.82\log(h_{Base}) - a(h_{Mobile}) + (44.9 - 6.55\log(h_{Base}))\log(d) + C_m$$

where:

$$a(h_{Mobile}) = (1.11\log(f) - 0.7)h_{Mobile} - (1.56\log(f) - 0.8) \qquad 1$$

and *Cm* = 0 dB, for medium sized city and suburban centres with medium tree density, while for metropolitan centres *Cm* = 3 dB. The model is restricted to the values of Table 1. The application of this model is restricted to large and small macro-cells, i.e., BS antenna heights above roof top levels adjacent to the BS.

Table 1: Model parameter range

Frequency	1500…..2000 MHz
h_{Mobile}	1…10 m
h_{Base}	30…200 m
Distance d	1…20 km

2.2 Link Budget

Link budget is used in calculating the maximum allowable propagation between two radio links. The Table 2 below shows an example of link budget in UMTS network. The term interference margin or noise rise (in row J) is responsible for the compromise between coverage and capacity. The higher the number of user (capacity) in the cell the higher the noise, this leads to increase in the interference and thereby decreases the cell range since the maximum allowable pathloss is reduced.

Table 2: An example of UMTS power budget for speech and data services.

	Parameter	Speech		Data		
		Downlink	*Uplink*	*Downlink*	*Uplink*	*Units*
A	Information rate	12.2	12.2	384	64	Kbps
B	Load	50	50	75	80	%

C	TX power	33	21	37	21	dBm
D	Cable/body loss	5	2	5	2	dB
E	TX antenna gain	18	0	18	0	Dbi
F	Peak EIRP	46	19	50	19	dBm
G	Thermal noise density	-173.93	-173.93	-173.93	-173.93	dB/Hz
H	Receiver noise figure	8	4	8	4	dB
I	Noise power at receiver	-100.13	-104.13	-100.13	-104.13	dBm
J	Interference margin	3.01	3.01	6.02	6.99	dB
K	Total noise power at receiver	-97.12	-101.12	-94.11	-97.14	dBm
L	Processing gain	24.98	24.98	10	17.78	dB
M	Required Eb/No	7	5	1.5	2.5	dB
N	Receiver sensitivity	-115.10	-121.10	-102.61	-112.42	dBm
O	RX antenna gain	0	18	0	18	dBi
P	Cable loss/body loss	2	5	2	5	dB
Q	Soft handover diversity gain	3	2	3	2	dB
R	Power control headroom Lognormal shadow margin	0	3	0	3	dB
S	Required signal level	-116.10	-133.10	-103.61	-124.42	dBm
T	**Allowable propagation loss**	**162.10**	**152.10**	**153.61**	**143.42**	**dBm**

3 Capacity Estimation

The approach used is the modified load equation. The load equation is commonly used to make semi-analytic estimation of the average capacity of a WCDMA cell without going into system level capacity simulation as described by [6]. There are a number of elements that influence the UMTS network load, which define the maximum number of users per cell. More over, the number of users and their bit rates has influence on the total throughput of a cell. Also the activity factor in speech and data services affects the load. However, the most important contributor to the load is the E_b / N_o requirement, which depends at least on the service type; the data rate of the service, the propagation conditions, and the receiver performance. The own-cell and other-cell interference also influences the load of a cell.

3.1 Uplink load equation

WCDMA radio interface offers enormous possibilities in obtaining large capacities; however, it imposes many limits as regards the acceptable level of interference in the frequency channel. In every cellular system with spreading spectrum, the radio interface capacity is drastically limited due to the occurrence of interference [2],

The relation E_b / N_0 for the j-th user as given by [7] can be calculated as follows:

$$\left(\frac{E_b}{N_o}\right) = \frac{W}{R_j \upsilon_j} \cdot \frac{P_j}{I_{Total} - P_j} \quad , \qquad 2$$

in which the following notation has been adopted:
P_j – received signal power from the j user,
W – chip rate of spreading signal,
v_j – activity factor of the j user,
R_j – bit rate of the j user,
I_{total} – total received wideband power including thermal noise power.
The mean power of the j-user can be determined with the help of the equation 3:

$$P_j = \frac{I_{total}}{1 + \frac{W}{\left(\frac{E_b}{N_o}\right)_j R_j \upsilon_j}} = L_j I_{total} \quad , \qquad 3$$

where L_j is the load factor for the user of jth connection.

$$L_j = \frac{1}{1 + \frac{W}{\left(\frac{E_b}{N_o}\right) R_j \upsilon_j}} \quad , \qquad 4$$

Once the load factor of a single user has been established, it is possible to determine the total load for the uplink connection:

$$\eta_{UL} = \sum_{j=1}^{N} L_j \ , \qquad 5$$

where, N is the number of users.

The above relation is true only when we deal with a system that consists of a single cell. In fact, however, there are many cells available, in which the generated traffic influences the capacity of radio interface of other cells. Thus, the relation 5 should be complemented with an element that would take into consideration interference coming from other cells [7]. To achieve this, a variable i is introduced, which is defined as other cell interference over own cell interference.

The total load for the uplink can thus be rewritten as:

$$\eta_{UL} = (i+1)\sum_{j=1}^{N} L_j = \sum_{j=1}^{N} \frac{1}{1+\frac{W}{\left(\frac{E_b}{N_o}\right)_j R_j \upsilon_j}}(i+1), \qquad 6$$

The bigger the load of a radio link, the higher level of the noise generated in the system. The increase in the noise Δ_{nr} is defined as the relation of the total noise received in the system to the thermal noise and is described by the following equation:

$$\Delta_{nr} = \frac{I_{total}}{P_N} = \frac{1}{1-\eta_{UL}}, \qquad 7$$

where P_N is the thermal noise.
The noise rise or interference margin is then given by:

$$IM_{UL} = 10\log_{10}\left(\frac{1}{1-\eta_{UL}}\right), \qquad 8$$

as described by [1].

When the load of the uplink approaches unity, the matching increase in noise tends towards infinity. Therefore, it is assumed that the actual maximal usage of the resources of a radio interface without lowering the level of the quality of service will amount to about 50 – 80% [2].

For the case of dimensioning, we assume a single service since UMTS services are expected to be demand driven. Therefore, equation 8 for a service class then becomes,

$$\eta_{UL} = \frac{N}{1+\frac{W}{\left(\frac{E_b}{N_o}\right) R\upsilon}}(i+1), \qquad 9$$

3.2 Modification of load equation (Imperfect power control)

The bit energy-to-noise density used here is the targeted value with the assumption that the base station and mobile station power control is perfect. Due to Multipath propagation, the variation in E_b / N_o is according to log-normal distribution with standard deviation in the order of 1.5 dB and 2.5 dB as described by [3]. Therefore in the case of imperfect power control, the constant targeted value of (E_b / N_o) for users with service g needs to be replaced by its expected value.

$$E[E_b / N_o] = (E_b / N_o)_t e^{\frac{(\beta\sigma_C)}{2}}, \tag{10}$$

Where $\beta = (\ln 10)/10$

Therefore the load equation can be modified as follows:

$$\eta_{UL} = \frac{N}{1 + \frac{W}{(E_b/N_o)_t e^{\frac{(\beta\sigma_C)}{2}} R\upsilon}}(i+1), \tag{11}$$

Also, the number of simultaneous users per service class is given by

$$N = \eta_{UL}\left(\frac{1}{i+1}\right)\left(1 + \frac{W}{\left(\frac{E_b}{N_o}\right)_t e^{(\beta\sigma_C)} R\upsilon}\right) \tag{12}$$

In terms of the noise rise,

$$N = \left(\frac{1}{1-\Delta_{nr}}\right)\left(\frac{1}{i+1}\right)\left(1 + \frac{W}{\left(\frac{E_b}{N_o}\right)_t e^{(\beta\sigma_C)} R\upsilon}\right) \tag{13}$$

Where Δ_{nr} is the noise rise.

3.3 Modelling of quality of service (Blocking probability)

This thesis employs the uplink blocking probability calculation method for cellular systems with WCDMA radio interface and infinite source population using modified Erlang model [8].
Since the number of simultaneous user given by the above equation has taken into consideration the effect of noise rise and imperfect power control, the resulting *N* gives the number of channels available for particular user traffic class. Therefore users will be blocked based on load since *N* depends on the load.

UMTS services are basically categorised into two namely; packet and circuit switched services. The circuit switched services are modelled by Erlang B formulae based on Busy Hour Lost Calls Cleared (LCC) and the packet switched services are modelled by Erlang C formulae based on Busy Hour Lost Call Delayed (LCD) as in [9].

Depending on the service type, for a given number of channels we calculate the offered load ρ from either Erlang B or Erlang C model.

3.4 Dimensioning in Mixed Multi-rate Environment

Here we calculated the number of users for a particular or any traffic mix. The input parameters to this model are Mean Holding time and QoS for circuit switched services and Average Download size, error rate and/or the access delay time for packet switched services.

From the Erlang capacity, the desired average throughput by a Node B can be calculated and is given by

$$D = \rho x R \, , \qquad 16$$

where ρ is the Erlang capacity for a given quality of service and R, is the service data rate

For the packet services, loss probability is given by:

$$p_c = \frac{\tau_d}{\tau_h} \, , \qquad 17$$

where τ_d is delay time or the response time of the packet service and
τ_h is the mean holding time for the service and is given by:

$$\tau_h = \frac{X}{R} \, , \qquad 18$$

where X is the average download within the busy hour

Also, $$CarriedTraffic = \frac{OfferedTraffic}{1 - p_c} \, , \qquad 19$$

Let the actual Busy Hour Traffic without the overhead for re-transmission for each service class be given by X_r;

For Circuit based services, the only input parameter is the τ_h and therefore,

Average Data rate per User per Busy Hour for r*th* service from Erlang is given by,

$$ADRU_r = \frac{R_r x \tau_{hr}}{3600} \, , \qquad 20$$

For Packet Services, it is given as

$$ADRU_r = \frac{X_r}{3600(1 - p_{cr})} \, , \qquad 21$$

Therefore Total Average User Data Rate for a user within the Busy Hour is given by:

$$\sum_{r=1}^{M} ADRU_r \qquad 22$$

where M is number of services provided by the UMTS network
Combined with equation (16), Total number of users N_u can be calculated;

$$N_u = \frac{D}{\sum_{r=1}^{M} ADRU_r}, \quad 23$$

D must be such that Min [D_r] < D < Max [D_r]

The complete flowchart is given by Figure 1 above.

5 Numerical Result

We first simulated the effect of imperfect power control on system capacity based on load equation using MATLAB. Here, the target E_b / N_o ranges between 2.5 and 10 dB and the standard deviation is considered for cases 0, 0.5 dB, 1.5 dB and 2.5 dB.

The result is plotted as shown in Figure 2 below. The graph indicated that the system capacity (number of simultaneous users) decreases with increase in standard deviation of the lognormal distribution of E_b / N_o as predicted in [3].
Through out the testing we use WCDMA network parameter as contained in Table3 and the same parameter is also for pathloss calculation.

Two bearer services 12.2 kbps voice service and 64 kbps data service are considered with E_b / N_o of 5 dB and 2.5 dB respectively, also with load factor of 50% and 80% respectively. This gives maximum allowable pathloss of *152.10 dB* and *143.42 dB* respectively which from Okumura propagation (medium urban) formula correspond to cell radius of *2.227 km* and *1.26 km* respectively.

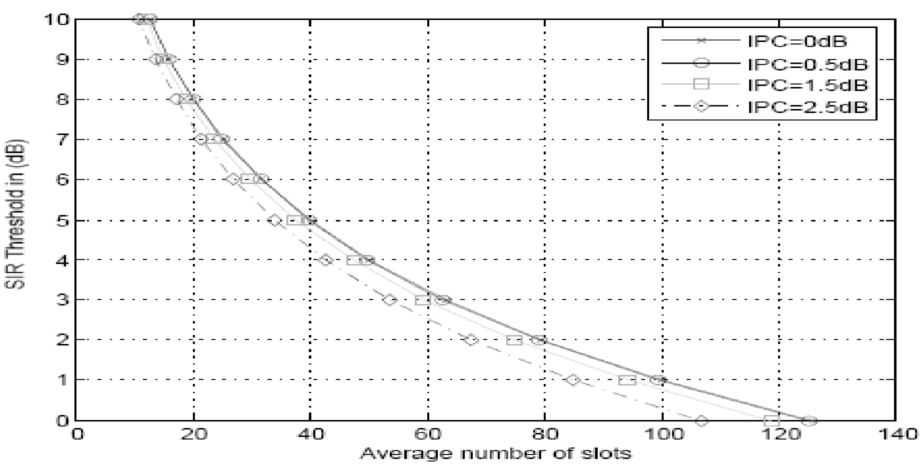


Figure2: Effect of Imperfect Power Control on Channel Capacity
This is obtained using the parameter in Table 3 below.

Table 3: Simulation parameter for the cell range

Frequency	2000 MHz
BS antenna height	30 m
MT antenna height	1.8 m

Cm	3 db

Then, from the load equation, the number of WCDMA channel corresponding to each load for particular service class is evaluated using the following parameter in Table 4.

Table 4: Simulation parameter for channel capacity

Parameter	**Voice (12.2 kbps)**	**Data (64 kbps)**
Eb/No	5 dB	2.5 dB
Power control deviation	1.5 dB	1.0 dB
Interference ratio	0.6	0.5
Load factor	0.5	0.8
Voice activity factor	0.6	1.0

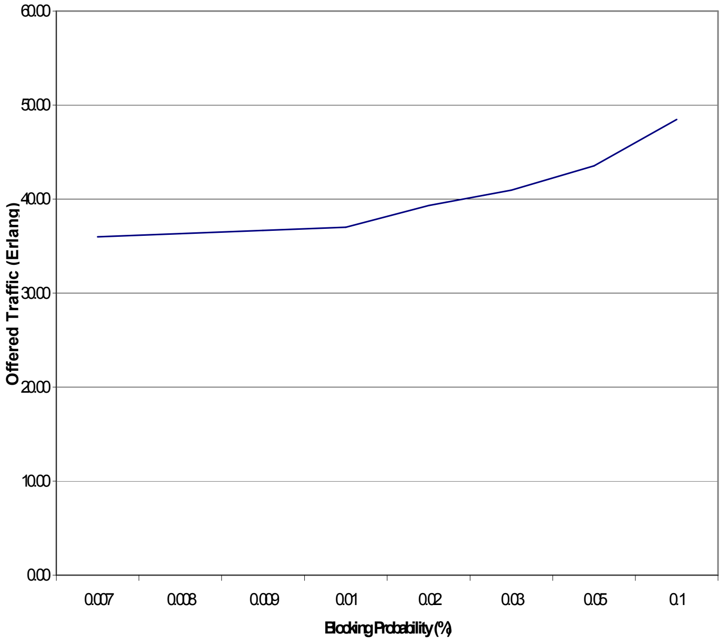
This yields approximately *49* channels for the 12.2 kbps voice service and *10* channels for the 64 kbps data service.

Figure 3 (a) and (b) show the relationship between offered load and the blocking probability for voice and data service respectively. We define the blocking probability as the probability that a user call will be dropped if channel is not available for the voice service and as the probability that a user will be delayed and queued because the channel is busy for the data service. Therefore, we assume that the voice service is circuit switched while the data service is packet switched. Consequently, we use Erlang B and Erlang C models respectively with the assumption that the connection request follows poisons distribution and can be model accordingly; also call sources are independent and the holding time is exponential.

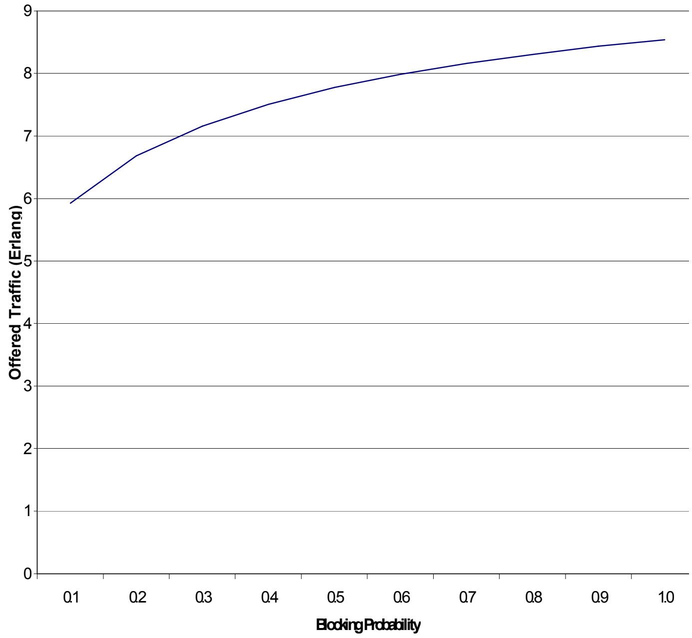
Also, for the mixed multi-rate environment, for example let an average user call for 90 s, download 10 Mb on 144 kbps link and 20 Mb on 384 kbps link, the following Table 5 contains the analysis using the above algorithm.

From the case study, average throughput from the Node B at loss probability of 0.02 according to equation 16 is approximately 480 kbps. In fact, it has been shown that depending in the parameters value, planned transceiver capacity is typically from 400 kbps to 700 kbps per transceiver [1][10]. Therefore, the number of subscriber that can use the transceiver is 51 for this particular traffic mix.

[1] Information obtained from http://www.umtsworld.com



(a)



(b)

Figure 3 (a): **Relationship between Offered Traffic and Blocking Probability for Voice service for N=49**
(b): **Offered traffic Vs Blocking Probability for 64 kbps Data service for N= 10**

Table 5: Analysis of mixed multi-rate Scenario

Service	Average data rate (kbps) per user per busy hour
Voice (12.2 kbps)	0.305
Data (144 kbps, $\tau_d = 5s$)	2.993
Data (384 kbps, $\tau_d = 5s$)	6.146
Total	9.384

6 Conclusion

The paper describes how to dimension UMTS radio interface analytically. Since Erlang model is robust in estimating the traffic, it is shown that the blocking probability still depends on E_b / N_o and the load factor which is determined by the noise rise in the cell. The paper shows that different traffic mixes can be combined together to dimension a transceiver or sector in the cell. Therefore the algorithm can be used for single service and mixed service. We are able to combine both coverage and capacity planning together and developed a practical approach in dimensioning UMTS network.

References

[1] Neimela J.: 'Impact of base station site and antenna configuration on capacity in WCDMA cellular network'. Master of Science thesis, Tampere University of Technology, Sweden, 2003.

[2] Laiho J.F., Jäsberg M., Sipilä K., Wacker A., Kangas A, "Comparison of three diversity handover algorithms by using measured propagation data", IEEE Proceedings Vehicular Technology Conference, pp. 1370 - 1374, Spring 1999.

[3] Viterbi, A.J. "CDMA Principles of Spread Spectrum Communication. Addison-Wesley", 1995.

[4] Hata,M., "Empirical Formula for Propagation Loss in Land Mobile Radio Services", *IEEE Trans. Veh. Technol.,* Vol. VT-29, No. 3, Aug. 1980, pp. 317-325.

[5] Okumura,Y., Ohmori,E., Kawano,T. and Fukuda,K., "Field Strength and Its Variability in VHF and UHF Land-Mobile Radio Service", *Rev. Elec. Commun. Lab.*, Vol. 16, Sep.-Oct. 1968, pp. 825-873.

[6] Glisic S., Vucetic B, "Spread Spectrum CDMA Systems for Wireless Communications", Artech House, 1997.

[7] Holma H. Toskala A., and Hämäläinen S., 1998. "Link and system level performance of multiuser detection CDMA uplink". Wireless Personal Communications 8, No. 3, pages 301-320. © 1998 by authors and © 1998 Kluwer Academic Publishers. By permission

[8] Erlang A.K., "On the rational determination of the number of circuits". In The life and works of A.K.Erlang. Brockmeyer E., Halstrom H.L. and Jensen A., eds. Copenhagen: The Copenhagen Telephone Company (1948).

[9] Freeman R. L. "Telecommunication System Engineering". 3rd Edition. John Wiley & Sons, Inc., New York, USA (1996).

[10] http://www.umtsworld.com/technology/capacity.htm "3G and UMTS Technology", Assessed 12th October, 2006.

ON THE POWER SPECTRAL DENSITY OF THE GSM SIGNALING SCHEME

V.O. MATTHEWS AND U.A.K. OKONKWO

Dept. of Electrical & Information Engineering, Covenant University, Ota
Anambra State University, Uli
vomatthews@yahoo.com , crucifi___xzo@yahoo.co.uk

Abstract

In this paper, the Power Spectral Density of encoded Gaussian Minimum Shift Keying (GMSK) which is the Signaling Scheme of the Global System for Mobile Communication (GSM) is derived by a combined approach of the autocorrelation method and Markov Process. In the analysis, the Amplitude Modulated Pulse decomposition proposed by P. Laurent is employed to ease computation. Encoding of the message data utilizes Convolutional Code of rate1/2. Results are for both the uncoded and coded waveform comparing variation in power spread over a range of frequency.

I Introduction

The ever increasing demand for digital wireless communication system presents a serious difficulty of spectral congestion that obviously causes severe adjacent and co channel interference problems. This has led to the investigations of a wide variety of techniques for solving the endemic problems that result from spectral congestion. Among the solutions to this problem are the use of frequency–reuse techniques, efficient source encoding techniques, spectral efficient modulation schemes and/or spectral efficient multiple access scheme.

The main objective of spectral efficient modulation format is to maximize the bandwidth usage at a prescribed bit error rate with minimum expenditure of signal (minimum signal power lay off). For GSM extra constraint is placed on the modulation scheme by the fact that non-linear amplifiers which operate near saturation are incorporated in the general architecture of the system. These nonlinear devices produce extraneous signal regrowth (sidebands) when passing a signal with amplitude fluctuations through them. So a modulation scheme for this must in addition be characterized by constant amplitude to combat such signal impairments. Thus at increased data rate, for low power consumption, and under the influence of nonlinear channel, we require a modulation format that balances the respective parameter requirements. This is a typical scenario in the GSM communication system.

A modulation technique which can offer this trade-off of complexity versus spectral efficiency is GMSK. This is a type of Continuous Phase Modulation (CPM), a class of nonlinear signaling scheme that are efficient in power and bandwidth. It also generates constant envelop waveform and therefore is very useful in radio channels employing non linear amplifiers like Traveling Wave Tube (TWT). GMSK has since been adopted as the modulation scheme for the GSM digital cellular system. The performance and analysis of CPM has been reported by several researchers. However, performance of encoding GMSK has not been examined fully. The analysis of coded some classes of CPM is reported in [2], [4], [7], [10], and [16]. It is obvious that channel encoding increases the required transmission bandwidth when considered independent of modulation and thus affect the power/bandwidth trade off of system. [1], [6], and [12] presented the Power spectrum analysis of uncoded CPM with some references to GMSK. The Power spectral density analysis of some type of encoded CPM is presented in [8], and [13].

In this paper, we present the Power Spectral Density of convolutionally encoded GMSK scheme. For ease of computation, the idea of decomposition of CPM signal presented in [3] is employed where the first two Laurent pulses [3] are used as the basic GMSK signal in power spectrum computation. The method of Spectral computation of Digital FM using Autocorrelation method is presented in [9]. In this report, the combine approach of autocorrelation/Markov method as discussed in [6] and [13] is used to model the encoding process of the GMSK signal in the course of power spectrum analysis.

II GMSK Representation

GMSK modulation is a modified form of Minimum Shift Keying (MSK), and a special case of binary CPFSK in which the modulation index h is set at 0.5. In this case, the rectangular shaping pulse used in conventional MSK is replaced by a special type of non linear pulse shaping filter called Gaussian filter. This scheme ensures a narrower spectrum than that of MSK [8]. Gaussian filter has an ideal impulse response given by

$$h(t) = K\sqrt{\frac{2\pi}{\ln(2)}}\,(BT_b)\exp\left(\frac{(-2(BT_b)\pi t)^2}{\ln(2)}\right)$$

(1.0)

For the case of GMSK analysis, this pulse can be modified thus

$$g_{GMSK}(t) = \frac{1}{2T_b}\left\{Q\left(\frac{2\pi B(t-\frac{T_b}{2})}{\sqrt{\ln(2)}}\right) - Q\left(\frac{2\pi B(t+\frac{T_b}{2})}{\sqrt{\ln(2)}}\right)\right\}$$

(2.0)

The primary parameter here is the (BT_b) product (-3db) of the Gaussian filter. In order to reduce the sidelobes and produce a compact spectrum, the appropriate value of (BT_b) should be used. In GSM the value of the BT parameter is 0.3. If the (BT_b) product is sufficiently large, then data sequence $\{a_K\}$ which often is None-Return-to-Zero (NRZ) will pass unfiltered. Smaller values of (BT_b) product will give a good compact spectrum where B is the half power bandwidth at symbol period T_b as shown in fig. 1.0

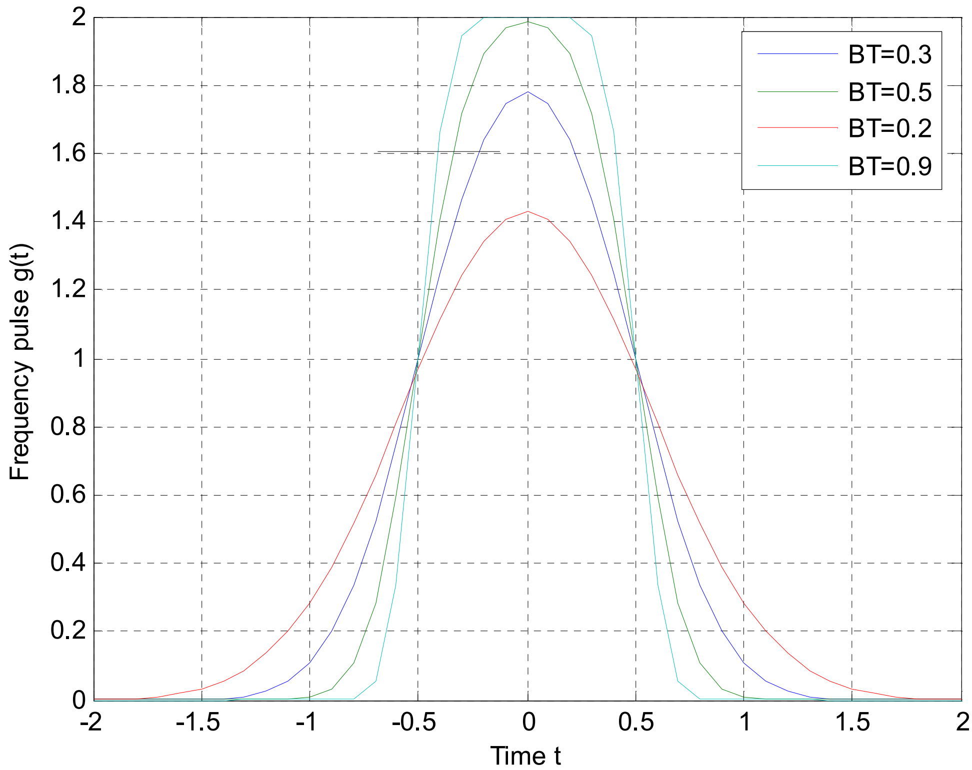


Figure 1: The frequency pulses of GMSK for BT= (0.2,0.3,0.5,0.9).

The GMSK waveform can be expressed as

$$S_{GMSK}(t) = \sqrt{\frac{2E}{T_b}}\, V\cos(2\pi f_c t \pm \pi h \sum_{k=1}^{\infty} a_K q_{GMSK}(t-kT_b))$$
$$= \sqrt{\frac{2E}{T_b}}\, V\cos(2\pi f_c t \pm \varphi(t,a)) \qquad (3.0)$$

where $\sqrt{\frac{2E}{T_b}}$ is the signal amplitude, f_c is the carrier frequency, a_k is the input data, T_b is the bit interval. The function $q_{GMSK}(t)$ is called the phase shaping pulse and is a continuous, monotonic function that determines the overall spectral characteristics of the modulated signal. It is defined as

$$q(t) = \int_{-\infty}^{t} g(\tau)\, d\tau = \frac{1}{2}\left[\frac{q_0\left(t-\frac{LT}{2}\right) - q_0\left(-\frac{LT}{2}\right)}{q_0\left(\frac{LT}{2}\right)}\right] \qquad (4.0)$$

where

$$q_0(t) = \frac{1}{4C}\left[A.erf(A) - B\,erf(B) + \frac{1}{\sqrt{\pi}}\exp(-A^2) - \frac{1}{\sqrt{\pi}}\exp(-B^2) + C\right]$$

$$C = BT\pi\sqrt{\frac{2}{\ln(2)}} \;;\quad A = C\left(\frac{t}{T}+0.5\right) \;;\quad B = C\left(\frac{t}{T}-0.5\right)$$

and $erf(x) = \frac{2}{\sqrt{\pi}}\int_0^{\infty}\exp(-y^2)\,dy$.

The information carrying phase is given by

$$\varphi(t,a) = \pi h\sum_{k=1}^{\infty} a_k q_{GMSK}(t-kT_b) \quad (5.0)$$

For $L = 3$ and $h = \frac{1}{2}$,

$$\varphi(t,a) = \pi\{a_k q(t-kT_b) + a_{k-1}q(t-(k-1)T_b) + a_{k-2}q(t-(k-2)T_b)\} + \frac{\pi}{2}\sum_{i=0}^{k-3} a_i \quad (6.0)$$

It can be shown that the unit amplitude complex envelope of (3.0) is of the form [11]

$$S_{GMSK}(t) = \varphi(t,a)\exp(j2\pi f_c t) \quad (7.0)$$

III Laurent Decomposition

The continuity imposed on the phase of GMSK signal depicts a kind of memory inherently built into it. The signal is best viewed as having a coded pattern directly imposed on the phase. This sought of encoding system employs an integrated approach to modulation and coding in which case, the encoding can take place in signal space as part of the modulation process. This approach offers the attractive possibility of achieving performance improvement without the bandwidth expansion which accompanies the usual concatenation of coding and modulation. The analysis of the inherent coding of CPM was shown to be achievable if one can decompose the CPM [3], [15], and [17].

In [3], P. Laurent showed that any constant amplitude binary phase modulation can be expressed as a sum of a finite number of time limited amplitude modulated pulses (AMP decomposition). This is the baseband signal that can be written as a sum of 2^{L-1} PAM signals expressed as

$$\tilde{S}_{GMSK} = \sum_{=0}^{2^{L-1}}\sum_{n=0}^{N-1} a_{k,n}C_k(t-nT_b) \quad (8.0)$$

where

$$C_k = S_0(t)\sum_{i=0}^{L-1} S_{i+a_{K,i,L}}(t) \quad (9.0)$$

$$S_n(t) = \frac{\sin(2h\pi\, q(t))}{\sin(\pi h)} \quad (10.0)$$

and the parameter $a_{k,i}$ is the message bit that takes on the value 0 or 1.

For binary GMSK for GSM application, L = 3, thus we have 4 distinct pulse shapes made up of 3-fold distinct products of the $S_n(t)$ corresponding to 2^{L-1}. Thus [2, eqn (11)],

$$\begin{aligned} C_0(t) &= S_0(t)S_1(t)S_2(t) \quad ; \quad 0 \le t \le 4T_b \\ C_1(t) &= S_0(t)S_2(t)S_4(t) \quad ; \quad 0 \le t \le 2T_b \\ C_2(t) &= S_0(t)S_1(t)S_5(t) \quad ; \quad 0 \le t \le T_b \\ C_3(t) &= S_0(t)S_4(t)S_5(t) \quad ; \quad 0 \le t \le T_b \end{aligned} \quad (11.0)$$

In [5], it was asserted that $C_0(t)$ and $C_1(t)$ are the most significant pulse durations and carry most of the signal energy. Assuming ergodic process, we can thus express (8.0) as

$$\widetilde{S}_{GMSK}(t) = \sum_{n=-\infty}^{\infty}\widetilde{a}_{0,n}C_0(t-nT_b) + \sum_{n=-\infty}^{\infty}\widetilde{a}_{1,n}C_1(t-nT_b) \qquad (12.0)$$

where $a_{0,n}$ and $a_{1,n}$ are equivalent complex data symbols.

The major advantage of this decomposition approach is that it allows us to study the coding operation independent of the modulation operation. The scheme decomposes the inherent nonlinear characteristic of the CPM into finite number of linear models. Once the memory is made explicit it becomes possible to design trellis and Convolutionally encoded system for CPM. Such decomposition tends to reduce the complexity associated with calculating the Power Spectral Density.

IV Convolutional Encoding

Convolutional Coding applied to the source sequence is design to add more bits to the final bit sequence. In general, the Convolutional code is characterized as having rate $R = \frac{k}{n}$ and constraint length K, where k and n are the number of inputs and outputs respectively, in the encoder. Fig: 2.0 shows a general model for convolutional encoding of GMSK, where $R = \frac{1}{2}$, and K=5.

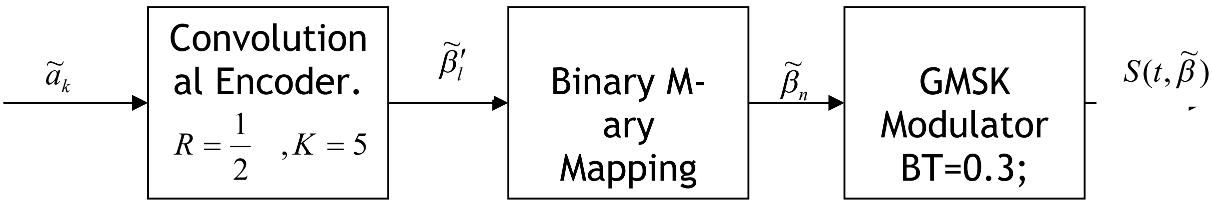


Figure 2.0: Block model of convolutionally encoded GMSK modulator.

The Convolutional polynomials for this system are respectively $G_1(X) = X^4 + X^3 + 1$ and $G_2(X) = X^4 + X^3 + X + 1$. In the figure above, a sequence of bits a_k of rate $\frac{1}{T_b}$ is the input to a rate $R = \frac{1}{2}$ convolutional encoder, producing 2 output code bits. For simplicity, we assume there are n bits mapped onto an M-ary modulation set where $M = 2^n$ or we deem it that the output of the encoder is serially fed into the modulator using an XOR gate. The basic action of encoding/mapping affects only the input into the GMSK modulator. The channel coding invariably affects both the bit interval and energy per bit. This is so since for the uncoded bit interval T_b, the coded bit interval T_c should be such that the same bandwidth is maintained, thus, $T_c = RT_b$. In the case of the energy per bit E, to maintain the same signal energy level, $E_c = RE_b$. So the energy of the uncoded bits is spread among the more numerous coded bits.

For the input complex data $\widetilde{a}_k = [\widetilde{a}_0, \widetilde{a}_1, \widetilde{a}_2 ...., \widetilde{a}_{k-1}]$, without encoding, it is expected that for $\widetilde{a}_i$ *in*, then $\widetilde{a}_i$, *out* *for* $\widetilde{a}_i = [0,1]$. There is virtually no effect on the input to the modulator. However, with the encoding process, the input to the modulator can be described thus

$$\widetilde{a}_i ,in \rightarrow [\widetilde{a}_{i,\theta_m}, \widetilde{a}_{i,\theta_m}],out \quad ; \quad for \quad i,m = [0,1,2,...] \qquad (13.0)$$

Hence we can express the output of the encoder as

$$\tilde{\beta}_n \quad = \quad [\tilde{a}_{i,\theta_m,0}, \tilde{a}_{i,\theta_m,1}, \tilde{a}_{i,\theta_m,2}, \ldots\ldots\ldots\ldots\ldots\ldots, \tilde{a}_{i,\theta_m,n-1}] \qquad (14.0)$$

θ_m depends on the state of the encoder memory. Thus, the state vector for the system can be expresses as

$$\Phi \quad = \quad (\tilde{\beta}_{n-1}, \tilde{\beta}_{n-2}) \qquad (15.0)$$

where $\tilde{\beta}_n$ is the transmitted bit at time *n*.

Thus eqn (12.0) becomes

$$\tilde{S}^e{}_{GMSK}(t) \quad = \quad \sum_{n=-\infty}^{\infty} \tilde{\beta}_{0,n} C_0(t - 2nT_c) + \sum_{n=-\infty}^{\infty} \tilde{\beta}_{1,n} C_1(t - 2nT_c) \qquad (16.0)$$

where $\tilde{\beta}_{0,n}$ and $\tilde{\beta}_{1,n}$ are equivalent encoded complex data symbols. The superscript e denotes encoded signal.

V Power Spectral Density

The encoded envelop signal has a generic expression of a complex signal with Real and Imaginary parts. The real part would be given as

$$\tilde{S}^e_{GMSK}(t) \quad = \quad \mathrm{Re}\{\varphi(t, \tilde{\beta}) \cos(-j2\pi f_c t)\} \qquad (17.0)$$

For simplicity, let $\varphi(t, \tilde{\beta}) \equiv \theta(t)$, then using Euler's identity,

$$S^e_{GMSK}(t) = \frac{\theta(t)\exp(j2\pi f_c t) + \theta(t)^* \exp(-j2\pi f_c t)}{2} \qquad (18.0)$$

The autocorrelation function is given by

$$R_{S_{GMSK}} R_{S_{GMSK}}(t, t+\tau) = E\left\{ \frac{\theta(t)\exp(j2\pi f_c t) + \theta^*(t)\exp(-j2\pi f_c t)}{2} \times \frac{\theta(t+\tau)\exp(j2\pi f_c(t+\tau)) + \theta^*(t+\tau)\exp(-j2\pi f_c(t+\tau))}{2} \right\}$$

$$\begin{aligned} &= \tilde{R}_{\theta\theta}(t, t+\tau)\exp(j2\pi f_c(2t+\tau)) + R_{\theta\theta^*}(t, t+\tau)\exp(-j2\pi f_c \tau) \\ &+ R_{\theta\theta^*}(t, t+\tau)\exp(j2\pi f_c \tau) + \tilde{R}_{\theta\theta^*}(t, t+\tau)\exp(-j2\pi f_c(2t+\tau)) \end{aligned} \qquad (19.0)$$

But

$$\varphi(t, \tilde{\beta}) = \sum_{n=-\infty}^{\infty} \tilde{\beta}_n q_{GMSK}(t - 2nT_c) \qquad (20.0)$$

Then for a random signal, the Hermitian symmetric property shows that

$$\tilde{R}_{\theta\theta}(t, t+\tau) = \sum_{n=-\infty} \sum_{n=-\infty} \tilde{\beta}_n \tilde{\beta}_m q_{GMSK}(t - 2nT_c)(t + \tau - 2mT_c) = 0 \qquad (21.0)$$

Thus

$$\varphi(t)\varphi(t+\tau) = \varphi^*(t)\varphi^*(t+\tau) = 0 \qquad (22.0)$$

Hence,

$$R_{S_{GMSK}\,S_{GMSK}}(t, t+\tau) = \frac{1}{4}\{\ R_{\theta^*\theta}(t, t+\tau)\exp(j2\pi f_c \tau) + R_{\theta\theta^*}(t, t+\tau)\exp(-j2\pi f_c \tau) \qquad (23.0)$$

It can easily be shown that the Fourier transform of the time average is given by

$$S^e{}_{S_{GMSK} S_{GMSK}} = \frac{1}{4}(S_{\theta^*\theta}(f - f_c) + S_{\theta\theta^*}(f + f_c)) \qquad (24.0)$$

Where $S_{\theta^*\theta}(\bullet)$ denotes the power spectrum of the equivalent encoded complex base band modulation. If we assume data symbol uncorrelation, then it can easily be deduced that the Power Spectrum is given by [13]

$$S^{e}{}_{\theta\theta}(f) = \frac{\sigma^2}{2T_c}\left|Q_i(f)\right|^2 + \left(\frac{\mu}{2T_c}\right)^2 \sum_{n=-\infty}^{\infty}\left|Q_i\left(\frac{n}{2T_c}\right)\right|^2 \delta\left(f - \frac{n}{2T_c}\right) - \frac{1}{T_c}\mathrm{Re}\left[Q_i(f)Q_i{}^*(f)\right]$$

(25.0)

where σ^2 and μ are the variance and mean of the stationary random process. $Q_i(t)$ is the Fourier transform of the Laurent pulses $C_i(t)$. If the effect of discrete spectrum which in itself carries no information is neglected due to the effect of differential encoding associated with the source coding, and it is assumed that no spectrum is created by the filter actions (advantage of CPM) then equation is reduced to

$$S^{e}{}_{\theta\theta}(f) = \frac{\sigma^2}{2T_c}\left|Q_i(f)\right|^2 = \frac{\sigma^2}{2T_c}\left|F\{\ C_i(t)\right|^2 \qquad (26.0)$$

In a situation where encoding process induces memory effect on the data symbol ensuring data correlation, the Fourier transform of the correlation among data symbols is related to the variance by [13]

$$\sigma^2 \ = \sum_{K=-\infty}^{\infty} R_{\tilde{\beta}^2} \exp(-j2\pi\, f_c t) \qquad (27.0)$$

Thus

$$S_{\theta\theta}(f) = \frac{1}{2T_c}\sum_{K=-\infty}^{\infty} R_{\tilde{\beta}^2} \exp(-j2\pi\, f_c t)\ \left|F\{\ C_i(t)\right|^2 \qquad (28.0)$$

$$S^{e}_{\theta\theta}(f) = \frac{1}{2T_c}\ \left|F\{\ C_i(t)\right|^2 S_{\tilde{\beta}_i^2}(f) \qquad (29.0)$$

By the nature of convolutional encoding process, the probability of observing any particular value in the sequence can be deemed to depend on the preceding values. In most practical cases, this dependency is often on the immediate preceding (Previous) sample. A procedure so described is called Markov Process. In using the Markov method, the convolutional encoder can be modeled as a Markov source characterized by a transition matrix

$$P_t^n \ = P_t = \begin{bmatrix} p_{11} & p_{12} & \cdots\cdots . & p_{1K} \\ p_{21} & p_{22} & \cdots\cdots . & p_{2K} \\ \cdot & \cdot & & \cdot \\ . & . & \cdots\cdots . & . \\ . & . & & . \\ p_{K11} & p_{K2} & \cdots\cdots . & p_{KK} \end{bmatrix} \qquad \forall\ n \geq 1 \qquad (30.0)$$

Then the autocorrelation function can be expressed in terms of the transition matrix and the matrix of the correlation between the basic baseband pulses. If we assume binary sequence and let $\tilde{\beta}_{i,n} \ = \ (+1,-1); \{i = 0,1\}$, then the correlation function can be given for the first Laurent pulse as

$$R_{\tilde{\beta}_0^2}(k) = (1-2P_t)^k \qquad (31.0)$$

And for the second Laurent pulse [6]

$$R_{\tilde{\beta}_1^2}(k) = \begin{cases} & l = 0 \\ -j(1-2P_t)^3 & l = 1 \\ [-j(1-2P_t)]^l & l \geq 2 \end{cases} \qquad (32.0)$$

Where the transitional probability P_t is given by

$$P_t = \begin{bmatrix} 1-p_0 & p_0 \\ p_0 & 1-p_0 \end{bmatrix} \quad (33.0)$$

From eqn (29.0)

$$S_{\tilde{\beta}_0^2}(f) = \sum_{K=-\infty}^{\infty} R_{\tilde{\beta}_i^2}(k)\exp(-j4\pi f_c kT_c) \quad (34.0)$$

Thus for the first Laurent pulse $C_0(t)$

$$S_{\theta\theta}^{00}(f;P_t) = \frac{1}{2T_c}\left|Q_0(f)\right|^2 \left\{ \frac{4P_t(1-P_t)}{2(1-2P_t)(1+\sin 4\pi fT_c)+4P_t^2} \right\} \quad (35.0)$$

And for the second pulse $C_1(t)$

$$S_{\theta\theta}^{11}(f;P_t) = \frac{1}{2T_c}\left|Q_1(f)\right|^2 4P_t(1-P_t)\left\{ \frac{1}{2(1-2P_t)(1+\sin 4\pi fT_c)+4P_t^2} - 2(1-2P_t)1+\sin 4\pi fT_c \right\} \quad (36.0)$$

The cross correlation between complex data stream induced by the memory effect that characterizes the convolutional encoding process is accounted for by the expression

$$S_{\theta\theta}^{10}(f;P_t) = \frac{1}{2T_c}\left|Q_0(f)Q_1{}^*(f)\right|^2 8P_t(1-P_t)\left\{ \frac{1}{2(1-2P_t)(1+\sin 4\pi fT_c)+4P_t^2} + 1 \right\}\cos 4\pi fT_c \quad (37.0)$$

Finally, the Power spectral density of the encoded GMSK is given by

$$S_{\theta\theta}^{e}(f;P_t) = S_{\theta\theta}^{00}(f;P_t) + S_{\theta\theta}^{11}(f;P_t) + S_{\theta\theta}^{10}(f;P_t) \quad (38.0)$$

VI Result and Discussion

The power spectral densities of both the uncoded and encoded GMSK are evaluated for BT=0.3 and P_t *=0.543* and simulated using MATLAB software program. We have demonstrated that for rate ½ and constraint length K=5 the spectrum of the coded system has side lobes that are smoother but elevated than that of the uncoded system.

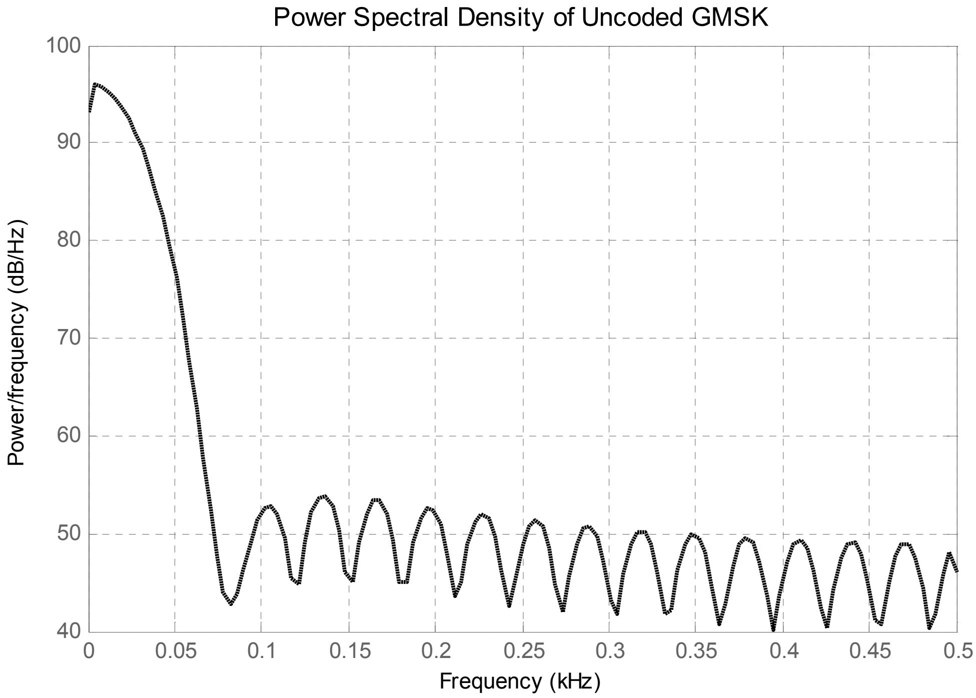


Figure 3.0: The Power Spectral Density of Uncoded GMSK

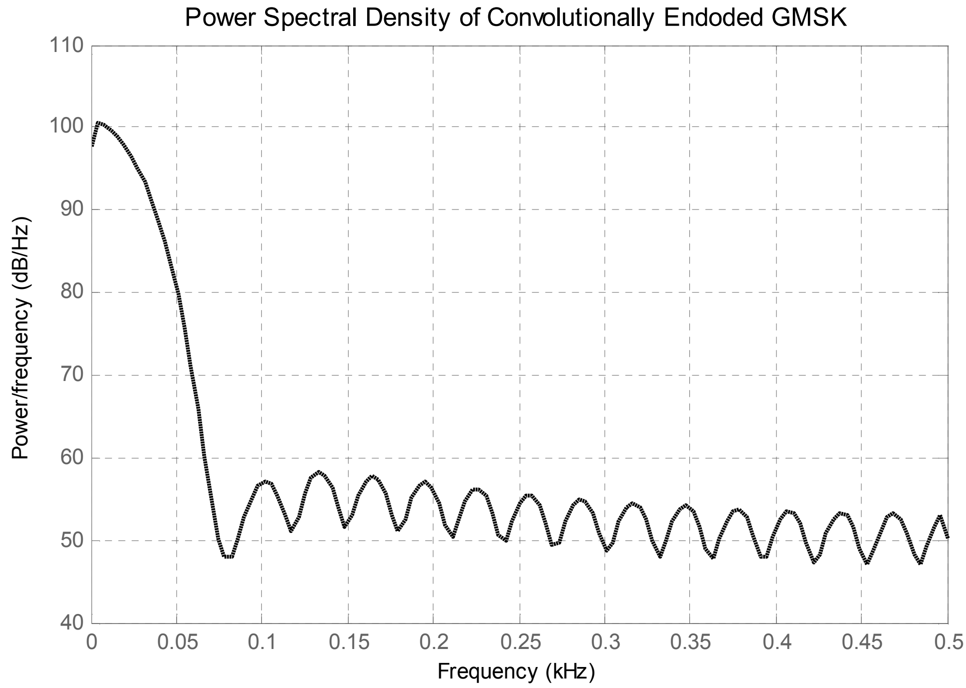


Figure 4.0: Power Spectral Density of Convolutionally Coded GMSK.

VII Conclusion

In this paper, we have derived the Power Spectral Density of Convolutionally Encoded Gaussian Minimum Shift Keying a Modulation technique adopted by the GSM Standard. The result of the MATLAB simulation of both uncoded and coded GMSK depicts variation in the Power Spectrum. . Peak-to-peak values measured from the average peak of the side-lobes to the peak of the main lobe for the uncoded and coded system are 41.91872dB and 42.39764dB respectively, showing a 0.47892dB deviation. The better performance of ½ rate encoded GMSK as compared to that of the uncoded system is due to the '*Spectral gain*' provided by the memory effect (inducing data correlation) of the convolutional encoding process employed, modeled as Markov process, which supposedly suppressed the spectral requirement from the added bit in the encoding process.

References

[1] K. Kuchi and V.K. Prabhu, "Power Spectral Density of GMSK Modulation using Matrix Methods*", MILCOM 1999-IEEE Military Com. Conference, No.1, Oct. 1999.*

[2] T. Svensson and A. Svensson, "On convolutionally Encoded Partial Response CPM",

[3] P.A. Laurent, "Exact and Approximate Construction of Digital Phase Modulations by Superposition of Amplitude Modulated Pulses (AMP)", *IEEE Trans. on Com., Vol. COM-34, No.2, Feb. 1986.*

[4] S.V. Pizzi and S.G. Wilson, "Convolutional Coding Combined with Continuous Phase Modulation", *IEEE Trans. on Com., Vol. COM-33, No. 33, Jan. 1985.*

[5] R. Booth, "A Note on the Application of the Laurent Decomposition to the GSM GMSK Signaling Waveform", www

[6] M.K. Simon, P.Arabshahi, L.Lam, and T.Y. Yan, "Power Spectrum of MSK-Type Modulations in the Presence of Data imbalance", *TMO Progress report 42-134, Aug. 15 1998.*

[7] D.K. Asano, and S. Pasupathy, "Efficient Block and Convolutional Coding of GMSK", *Proc. Of the 10th Inter. Symposium on Pers., Indoor & Mobile Radio Com. (PIMRC 1999).*

[8] J.B. Anderson, C.E. Sundburg, T. Aulin, and N. Rydbeok, " Power Bandwidth Preference of Smoothed Phase Modulation Codes*", IEEE Trans. on Com. Mar. 1981.*

[9] T. Aulin, and C.E. Sundburg, "Calculating Digital FM Spectral by means of Autocorrelation*", IEEE Trans. on Com., Vol. 30, No.5, May 1982.*

[10] S. Pasupathy, "Minimum Shift Keying: A Special Efficient Modulation", *IEEE Com. Mag. July 1979.*

[11] I. Korn, "Error Probability and Bandwidth of Digital Modulation", *IEEE Trans. on Com., Feb. 1980.*

[12] D. Divsalar, and M.K. Simon, " The Power Spectral Density of Digital Modulation Transmitted over Nonlinear Channels", *IEEE Trans. on Com., Jan. 1982*.

[13] U.A.K. Okonkwo, "Power Spectral Density of an Encoded Continuous Phase Modulation*", M.Sc. Thesis, University of Lagos, Nigeria, Aug. 2004.*

[14] L.H. Dong, "On the Source Matching Approach for Markov Sources*", IEEE Trans. on Information Theory, Vol. IT-29, No.5, Sept. 1983.*

[15] U. Mengali, and M. Morelli, "Decomposition of M-ary CPM Signals into PAM Waveforms", *IEEE Trans. on Information Theory, Vol. 41, No.5, Sept. 1995.*

[16] S.A. Gronemeyer, and A.L. McBridge, "MSK and Offset QPSK Modulation*", IEEE Trans. on Com., Vol. COM-24, No.8, Aug.1976.*

[17] B.E. Rimoldi, "A Decomposition Approach to CPM", *IEEE Trans. on Information Theory, Vol.34, No.2, Mar. 1988.*

IMPACT OF MOBILE AND WIRELESS TECHNOLOGY ON HEALTHCARE DELIVERY SERVICES

Z.O. OMOGBADEGUN AND C.K. AYO

Department of Computer & Information Sciences
College of Science & Technology, Covenant Universit, Ota, Nigeria
email: oniomogbadegun@yahoo.com[1], ckayome@yahoo.com[2]

Abstract

Modern healthcare delivery services embrace the use of leading edge technologies and new scientific discoveries to enable better cures for diseases and better means to enable early detection of most life-threatening diseases. The healthcare industry is finding itself in a state of turbulence and flux. The major innovations lie with the use of information technologies and particularly, the adoption of mobile and wireless applications in healthcare delivery [1]. Wireless devices are becoming increasingly popular across the healthcare field, enabling caregivers to review patient records and test results, enter diagnosis information during patient visits and consult drug formularies, all without the need for a wired network connection [2]. A pioneering medical-grade, wireless infrastructure supports complete mobility throughout the full continuum of healthcare delivery. It facilitates the accurate collection and the immediate dissemination of patient information to physicians and other healthcare care professionals at the time of clinical decision-making, thereby ensuring timely, safe, and effective patient care. This paper investigates the wireless technologies that can be used for medical applications, and the effectiveness of such wireless solutions in a healthcare environment. It discusses challenges encountered; and concludes by providing recommendations on policies and standards for the use of such technologies within hospitals.

Keywords: healthcare, mobile/wireless solutions, mobile architecture, healthcare portal, 3G Medicine

1. INTRODUCTION

Healthcare is prominent throughout the world. It represents a significant percentage of many countries' GDP and is growing steadily. But the increasing demands on healthcare systems are too often accompanied by shortages in clinical staff. Mobile technology and WLAN solutions promise to transform the healthcare industry. Patient-centered care, e-pharmacy, asset tracking, mobile voice and rich media systems are just a few of the solutions that are enabled by WLAN technology. Healthcare is one of the few verticals aggressively developing and extending applications and business processes to the WLAN, making the industry a hotbed of WLAN innovation [3]. In one perspective, the significant advancement and proliferation of wireless data technologies create entirely new areas of applications. Of particular importance is the potential impact of next generation wireless technologies on healthcare delivery. The major challenges of healthcare management in the 21st century include the commonalties of a global and apparent terminal malady of exponentially increasing costs, an informed and empowered consumer, the need for e-healthcare adaptability, and a shift from focusing on primarily curing to the prevention of diseases [1]. These technologies provide a variety of opportunities to address public healthcare challenges such as universal access for the uneducated, counter-bioterrorism, telemedicine, distance education, and home care. These opportunities present new challenges such as: surveillance, privacy/confidentiality/security of personal information which will affect all of our lives [4].One of the advanced information technology solutions gaining popularity is the wireless access to patient records and other healthcare services. This really is the third generation of wireless in healthcare. Doctors and other healthcare professionals have progressed from pagers to basic cell phones and now to data-enabled "smartphones". Physicians have been trading in their old, disconnected PDAs for modern PDA smartphones with cellular integration and built-in Wi-Fi and Bluetooth wireless functionality for use within the clinical environment [5].

It has become necessary to evaluate performance of protocols for wireless medical devices in clinical environments to achieve improvement in the safety and efficiency of medical device wireless communication and operation. Nowadays healthcare providers such as doctors and nurses can directly communicate with one another via wireless infrared beaming in their PDAs. In particular, modern healthcare institutions can benefit from more extensive use of mobile devices for pervasive and ubiquitous access to healthcare information systems to deliver healthcare services, such as general surgery and emergency services [6]. With a large mobile population of doctors, nurses, physician’s assistants and other caregivers, wireless LANs bring the ability to access the latest patient charts, medical records and clinical decision-support data at all times, anywhere in the healthcare organization. And as caregivers travel among different facilities, wireless network allows for easy connectivity at each site [7].

2.0 TECHNOLOGY AND HEALTHCARE SERVICES

2.1 Handheld Devices, Wireless Communications, and Real-Time Clinical Information Systems

Handheld portable computing devices promise to be strong catalysts in bringing the full power of information technology to the **point-of-care**. Personal digital assistant-type devices can significantly increase caregiver productivity and efficiency while preventing errors. These devices can be supported by wireless network and/or speech recognition Web-based and/or electronic data interchange (EDI) connectivity with all major suppliers to provide a more natural, easy-to-use interface for the clinician. In addition, they combine multiple functions (address book, scheduler, note pad, email, calculators, scanners, etc.) thus eliminating the need for multiple devices. A physician could access his or her patient schedule; document the clinical visit; order laboratory test, imaging, and medications; and generate data for billing purposes all from a single handheld wireless interface device. In the nearest future, the availability of ubiquitous, reliable wireless connectivity will provide anytime, anywhere access to information and messaging through highly portable or even wearable systems with speech interfaces and natural language command and control. In addition, cell phone-like devices for accessing the Web are poised to revolutionize Internet use. The use of point-of-care decision support will leapfrog physicians’ and hospitals’ efforts to improve patients’ safety and reduce the occurrence of medical errors [8]. Mobile devices are becoming increasingly popular across the healthcare field, enabling caregivers to review patient records and test results, access charge captures, enter diagnosis information during patient visits and consult drug formularies. Reduced medical errors, elimination of duplicate entries, increased accuracy of data, improved patient care, provision of the most up-to-date patient information available and the ability to update this information, and decreased operating costs are some of the benefits [2].

2.2. Mobile telemedicine

Pervasive healthcare has the potential to reduce long-term costs and improve quality of service, but it also faces many technical and administrative obstacles. Unique capabilities of emerging mobile device, wireless networks and middleware technologies can support a wide range of applications and services including mobile telemedicine, patient monitoring, location-based medical services, emergency response and management, pervasive access to medical data, personalized monitoring, and lifestyle incentive management.

2.3. Patient monitoring

Wireless LANs and personal area networks make it possible to continually monitor patients almost anywhere and immediately notify healthcare workers, the nearest hospital, or an emergency service of any critical change in status. Such networks can quickly route biomedical and environmental data from sensors deployed on the body, in a room, or throughout a building to a central computer system for processing [9]. Wireless technology enables ambulance personnel to send real-time data about a patient's condition to a hospital while en route.

2.4. Maturing Telemedicine Technologies

Maturing telemedicine technologies, struggling mobile networking revenues and increased personal healthcare awareness have provided the foundations for a new market niche for '3G Medicine'. During the last 5 years, telemedicine (based on internet and web technologies) is becoming a reality both in terms of developing technologies and supportive legislation. Within Europe, wireless infrastructures (3G Networking) has received a huge investment and although not well defined on how it would be achieved, healthcare has been identified as a major stream of revenue with personal healthcare being a key issue especially for the handset manufacturers. Combined with an increased awareness, not only for outpatients but also for the "well-worried" (healthy and health conscious), 3G Medicine Services will play an important role in personal healthcare management [10]. Wireless telemedicine is a new and evolving area in telemedical and telecare systems. Healthcare personnel require realtime access to accurate patient data, including clinical histories, treatments, medication, tests, laboratory results and insurance information. With large-scale wireless networks and mobile computing solutions, such as cellular 3G, Wi-Fi mesh and WiMAX, healthcare personnel can tap into vital information anywhere and any time within the healthcare networks. The recent introduction of pervasive computing, consisting of radio frequency identification (RFID), Bluetooth, ZigBee, and wireless sensor networks, further extends the potential for exploitation of wireless telecommunications and its integration into new mobile healthcare delivery systems [11].

3. ADOPTING WIRELESS SOLUTIONS IN HOSPITALS

3.1 Wireless Technology and Medical Applications

Healthcare providers look up to wireless solutions for a variety of reasons. Among the desired goals are reduced risk when monitoring patients, better patient outcomes, increased staff efficiency and continuity of care, improved response time, decreased costs associated with the care process, and compliance with government regulations and standards. Another goal is to reduce the cost and complexity of network management by leveraging existing wireless infrastructures and by deploying clinical and other value-added applications [12]. Point-of-care applications offer tremendous potential for today's healthcare industry as they enable healthcare providers to deliver service when and where it is needed, while still providing timely and secure access to all the critical information required. A wide variety of healthcare applications with industry-leading mobile and wireless technologies provide solutions such as:

- **Mobile Care Delivery:** Record patient information at the point-of-care; Gather patient history; Monitor vital signs; and Ambulatory car.
- **Physician's Orders and Results:** Prescription writing; Drug administration monitoring; Drug interaction monitoring; Laboratory specimen collection; Track samples from bedside to lab; Review test results; and Charge capture.
- **Intelligent Devices:** Diagnostic Devices; Kidney dialysis, anesthesia delivery; Patient monitoring; Blood glucose monitoring; Medication dispensing devices; and Home disease management.
- **Sales Force Automation:** Pharmaceutical sales; and Medical equipment sales.

Bluetooth technology has become a wireless technology well suited to use in battery-powered commercial electronic devices, such as PDAs, cell phones, and a variety of other mass-produced electronics [9].

3.2. Wireless Patient Monitoring

As hospitals deploy mobile applications to improve operational efficiency, wireless networks will be installed initially for high patient throughput areas, including emergency rooms, critical care wards and nursing care floors. Today, the adoption of wireless standards such as IEEE 802.11b/g (also known as Wireless Fidelity or "Wi-Fi"), and the use of mobile computing platforms form the basis of this wireless revolution. At the same time, clinical caregivers want to do more on a mobile basis and have grown accustomed to receiving information in real time via cell phones, pagers and PDAs. These factors have converged and are providing the impetus for a wide acceptance of wireless technologies within the healthcare industry [13].

4.0. CRITICAL WIRELESS INFRASTRUCTURE'S VULNERABILITIES

4.1. Access Point (AP) Vulnerabilities and Mobile Trust Model

Software that runs critical wireless infrastructure can suffer from vulnerabilities just like any complex system - often with devastating effects. A single vulnerability may compromise a host or a group of hosts, but a vulnerability in an infrastructure device can lead to much more damaging effects. Many access point (AP) vulnerabilities are caused by problems with the access points' administrative interfaces. A great way to protect yourself from vulnerabilities that might crop up is to disable access to APs on the wireless side and employ packet filtering on the wired side to limit access to a few select management workstations[14]. The movement towards Web-based services, and the increasing dependency on the Web have also made reliability a first-rate security concern. From malware and spyware, drive-by downloads, typo squatting, denial-of-service attacks, to phishing and identity theft, a variety of threats make the Web an increasingly hostile and dangerous environment. By undermining user trust, these problems are hampering commerce and the growth of online communities [15]. Given that systems will be transmitting highly sensitive information namely patient data, implicit in their use is a need for high level of end-to-end security, confidentiality and privacy. Mobile trust model can be incorporated into any wireless healthcare initiative that will enable healthcare organizations to meet the necessary security standards. By utilizing a well designed trust model in the structuring of a mobile or wireless initiative, it will only then be possible for these organizations to maximize the benefits from wireless technologies as well as minimizing their risks; thereby, enabling the benefits of cost- effective, quality-healthcare to those who are most critical; namely the patient [16].

4.2. Wireless sensor networks in healthcare applications

With the proliferation of handheld and mobile computing devices, it is important for management and information security leaders to be aware of the overriding security issues that accompany the use of these devices, and to ensure the users of these devices use them in a secure, and approved, manner [17]. Moving to a fully pervasive system would be a complex transition requiring several steps and incremental budgetary increases to create the necessary infrastructure [9]. The flexibility and richness of the Web architecture have come at the price of increasing complexity and lack of sound overall security architecture. The emergence of wireless sensor networks (WSNs) in healthcare applications is gaining momentum through the increasing array of wearable vital sign sensors and location tags which can track both healthcare personnel and patient status/ location continuously in real-time mode. Generally, WSN devices are extremely limited in terms of power, computation, and communication. They are often deployed in accessible areas, thus increasing security vulnerabilities. The dynamic ad hoc topology, multicast transmission, location awareness, critical data prioritisation, and co-ordination of diverse sensors of healthcare applications further exacerbate the security challenges [18].

5.0. EFFECTIVENESS OF WIRELESS HEALTHCARE SOLUTIONS

5.1. Security with Trust

The effectiveness of wireless healthcare solutions will be a function of security with trust and the appropriateness of the wireless trust model to a particular healthcare environment. Mobile healthcare needs security. In other words, wireless trust environment closely relates to individual intention to adopt wireless technology. Mobile healthcare security embraces confidentiality, authentication, and message integrity, and must also be seen in the broader context of e-healthcare systems. Since mobile devices can be lost or stolen more easily than their fixed counterparts and a guaranteed physical protection of mobile devices such as PDAs is not very practical, a systemic solution is needed to establish identity of the provider even if the provider is using a preauthorized device. A provider who is interested in engaging in a trusted transaction needs to meet two tests irrespective of access location: (1) The provider must be able to identify his or her identity; and (2) The provider must be able to demonstrate the ownership of the mobile device being used for the transaction. Mobile devices carry hardware-based Subscriber Identification Modules (SIM) that can be used to authenticate the device [16]. End-to-end security of a mobile transaction requires intervention at multiple points since none of the individual technologies and devices offer a complete security solution.

5.2. Protecting patient confidentiality

While the promise of better healthcare through better access to medical records is beguiling, it also creates a new layer of responsibility for physicians. When it comes to technology, it is imperative that physicians use only those devices that employ the very highest standards of security available. Privacy and security are also potential problems. Healthcare data should be available anytime anywhere, but only to authorized persons. Confidentiality in a mobile transaction is possible if the content of the transaction can be protected through end-to-end encryption. Not only the content must be encrypted in the mobile device itself, it must be transmitted in an encrypted form to the hospital's application gateway and vice versa. This requirement should not be confused with the radio path encryption since the content of a transaction uses media beyond the radio path [19]. Pervasive healthcare information could be abused by corporations in deciding who should be promoted, by insurance companies in refusing coverage for people with poor health, and by spouses and their attorneys in divorce cases. The large-scale introduction of wireless technology in healthcare has legal and regulatory implications [9].

5.3 PROPOSED SYSTEM FOR HEALTHCARE SERVICES

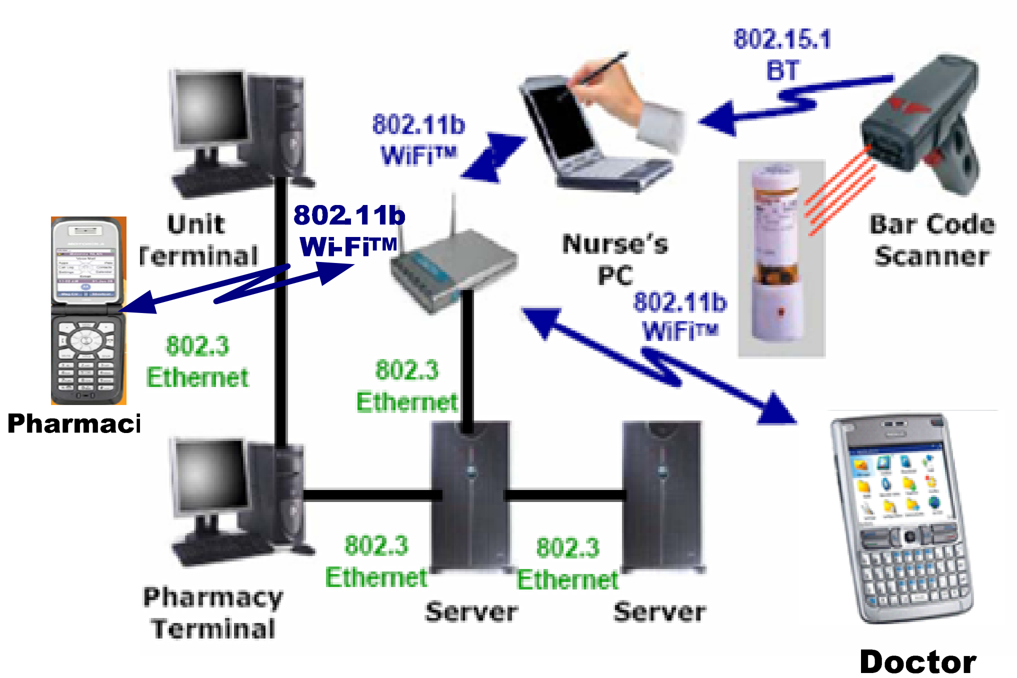


Figure 1 Wired/Wireless System for Healthcare Delivery Services

The IEEE standard 802.16, also known as Worldwide Interoperability for Microwave Access (WiMAX), is viewed as a powerful enabler for the massive deployment of high performance metropolitan area networks. Medication errors are the most significant and pervasive result of patient mis-identification. Poor patient identification can seriously affect the health of patients, as well as hospital bottom line. Patient identification problems in hospitals are particularly distressing because most of them do not have to happen. Using bar codes or RFID technology for positive patient identification is the fastest, easiest way to ensure accuracy in labeling drugs and patient samples, preparation and distribution of prescriptions, and administering correct medication dosages. The 802.11b standard is the first standard to make WLANs usable in the general workplace by providing robust and reliable 11 Mbps performance, five times faster than the original standard. Pharmacists can send prescription requests without the need to return to the dispensing department and doctors can sign patient prescriptions remotely. Surgeons can plan operation schedules and resources directly as a result of a consultation, allowing them to spend more time on operations, rather than administration. From the theatre itself, a timetable can be generated, and the required nurses, anaesthetists, and equipment sourced instantly, something that traditionally requires a lot of phone calls and meetings to co-ordinate [20].

6.0 CONCLUSION

Healthcare requires a new way to manage current challenges and to provide cost-effective quality care to all. Wireless technologies can reduce medical costs and improve quality of service. Provision of seamless mobility and connectivity to patients and frontline caregivers across the continuum of care through wireless patient wearable vital signs monitors, productivity tools, IP phones, PDAs, seamless enterprise coverage, robust fault tolerant WLANs, and integration of biomedical and IT systems on a single network infrastructure, conventional Telemetry and New Telemetry Offerings, Communications systems (Paging, In-building phones, Asset tracking), and Standards: 802.11 constitutes the Wireless Networks Solution offering the required comfort on the impact of mobile and wireless technologies on healthcare services delivery.

In this paper, numerous benefits have been highlighted with the adoption of wireless technologies in healthcare delivery. The adoption of wireless standards (e.g. IEEE 802.11b/g); use of mobile computing platforms; access to information in realtime via cell phones, pagers, and PDAs through availability of ubiquitous, reliable wireless connectivity; have demonstrated the suitability and wide acceptance of wireless technologies in healthcare. One major obstacle in mobile adoption of and development in healthcare is trust as there is a need for high level of end-to-end security, confidentiality, and privacy. The security gap existing between wireless sensor networks (WSN) and the requirements of medical applications needs to be resolved. On-demand encryption of wireless and communication protocols used to connect a mobile device to a base station offers a partial solution. A comprehensive security strategy that guarantees confidentiality, authenticity, content integrity, and nonrepudiability must be put in place to combat the vulnerabilities of wireless technologies in healthcare.

References

[1] Steve Goldberg and Nilmini Wickramasinghe (2006): (http://csdl2.computer.org/comp/proceedings/hicss/2003/1874/06/187460163b.pdf

[2] Solutions for Healthcare(2006): http://www.ianywhere.com/industries/healthcare_brochure.html.

[3] Siemens (2005): Wireless LAN in Healthcare**,** http://www.vocera.com/downloads/Siemens_Healthcare_Fall_2006v6_3.pdf:

[4] Luis G. Kun (2001): Telehealth and the global health network in the 21st century, From homecare to public health informatics, *Computer Methods and Programs in Biomedicine*, Volume 64, Issue 3, Pages 155 - 167 (01 March 2001)

[5] Neil Versel (2003) Use of Mobile and Wireless Technology Jumps in Hospitals http://tmlr.net/jump/?c=23738&a=296&m=4292&p=1136891&t=164

[6] Patrick C. K. Hung (2006): Research Issues of Healthcare Privacy Framework in Mobile Ad Hoc Networks (MANETs), Faculty of Business and Information Technology University of Ontario Institute of Technology (UOIT), Canada http://www3.uninsubria.it/uninsubria/allegati/pagine/2271/MANET-PatrickHung-Abstract12.PDF

[7] Meru (2005) Application Note: Wireless LANs and Healthcare: Understanding Security to Ensure Compliance with HIPAA, http://www.merunetworks.com/pdf/AP4-0805_Wireless_LANs&Healthcare.pdf

[8] Golembesky, Henry E (2001): Technology Approaches That Are Changing the Provider Health Care World and Transforming Performance, Group Practice Journal, 50(3), pp11-20, March 2001, The American Medical Group Association

[9] Upkar Varshney [2006], Pervasive Healthcare C O M M U N I C A T I O N S, Department of Computer Information Systems at Georgia State University.

[10] A. Marsh (2002). 3G Medicine - The Integration of Technologies, Computational Science - ICCS 2002: International Conference, Amsterdam, The Netherlands, April 21-24, 2002. Proceedings, Part II, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Volume 2330/2002, p. 972

[11] H. S. Ng, M. L. Sim, C. M. Tan and C. C. Wong (2006) Wireless technologies for telemedicine, BT Technology Journal, Springer Netherlands, Volume 24, Number 2 / April, 2006, pp. 130-137

[12] Kennedy, S (2006): Challenges to Implementing Wireless LANS in Healthcare, http://www.atmeda.org/news/2002_presentations/kennedy.challenges.ppt

[13] Dräger Medical AG & Co. KgaA (2005): Special Projects of National Significance 2005 Report to CARE Act Grantees: Wireless Patient Monitoring: Part II: Technical Perspective, PDF1112 / 06.05-2 / dp / Subject to modifications /, Published by Dräger Medical, September 2005.

[14] Chris Waters (2006): Keeping A Wireless Network Safe From Product Vulnerabilities: *Wireless Security,Networkworld.Com, 04/10/06, Http://www.networkworld.com/columnists/2006/041006-wireless-security.html*

[15] www2007.CFP (2007): 16th International World Wide Web Conference, Call for Papers Brochure, Banff, Alberta, Canada, May 8-12, 2007. Accessed Nov 24, 2006

[16] Binshan Lin (2006): Trust in Mobile Healthcare: Research Review and Implications, Louisiana State University in Shreveport,http://www.iacis.org/iis/2003_iis/PDFfiles/Lin.pdf

[17] CSI (2002): 12th Annual NetSec2002 Technical Dimensions of Network Security Conference Brochure, San Francisco, USA, June 17-19, 2002, p. 14

[18] H. S. Ng, M. L. Sim and C. M. Tan (2006): Security issues of wireless sensor networks in healthcare applications, BT Technology Journal, Springer Netherlands , Volume 24, Number 2 / April, 2006, pp 138-144

[19] Siemens (2006): Wireless Security and Healthcare – Going Beyond IEEE 802.11i to Truly Ensure HIPAA Compliance http://wp.bitpipe.com/resource/org_959010341_819/11693_35664_edp.pdf?site_cd=bp

[20] Sarah Payne (2003): Faith, hope and charity, http://www.bluesocket.com/pdf/health/GoodHopeHospital-WirelessCaseStudy.pdf

THE PROSPECTS OF M-VOTING IMPLEMENTATION IN NIGERIA

U.O. EKONG AND C.K. AYO

Department of Computer and Information Sciences,
Covenant University, Ota, Ogun State
vekong1@yahoo.com, ckayome@yahoo.com

ABSTRACT

Since independence, an average of 50% of registered voters participates in voting [1]. Similarly, an increasing rate of apathy was observed between the electorate and the elect, which was not unconnected with lack of transparency, accountability, and probity on the part of government [2]. Thus the electorate did not see the need to subject itself to any stress. Consequently, government is very committed to implementing the forth coming elections through e-voting. This paper proposes the prospects of m-voting implementation in Nigeria through the use of mobile phones, PDAs, etc. with guaranteed security, secrecy, and convenience in a democratization process. It also reviews the level of adoption of GSM in Nigeria, the implication of voting through the GSM, and finally introducing m-voting innovation in the voting process to increase voters' access and participation rate in elections.

Keywords: M-voting, Mobile devices, M-government, Mobile application and E-voting.

1.0 INTRODUCTION

According to Douglas Jones in [3], the first paper ballot began to replace oral voting in Rome in 139BC, and in the United States beginning in 1629. The first modern paper ballot, called the Australian ballot, was used in Australia in 1959. As of 1996, paper ballots were still used by 1.7% of the registered voters in United States. This method of voting is what the Nigerian system has been using in four out of five past elections that is, in 1979, 1983, 1999 and 2003 [4]. Countries in Georgia were the first jurisdictions to use punch cards and computer tally machines [5]. The marksense ballot, which utilized optical-scan technology, was used by 24.6% of registered voters in the United States for the 1996 presidential election, and their use is on the rise. Direct Recording Electronic Voting Machines (DRE) is a microcomputer implementation of a lever machine, invented in 1978. According to Kevin Bensor in [5], voters directly enter choices into electronic storage with the use of a touch screen, push buttons, or similar devices.

Electronic voting system (E-Voting) is a term encompassing several different types of voting. Electronic voting can include kiosks, the internet, telephones, punch cards, and marksense or optical scan ballots. Rick Semiatin, an American University Professor remarked that the optimal voting system is an electronic one with redundancy, or automatic back-ups, built into it [6]. Several states in the United States ran pilot programs on the Presidential Election Day 2000 to test e-voting as a valid process. These states include Arizona and California [7]. Other countries that have tested and found e-voting acceptable include Brazil, Belgium, Canada, India, Ireland, Geneva Venezuela, [4]. In California, which has a very good vote counting system, lot of it is electronic. This contradicts the punch card system used in Florida. The advantages of e-voting include streamlining the voting process, preventing ballot errors and confusion and increasing national voters turnout [5]. Computers would streamline the voting process because they do not use paper ballot or bulky machinery. Computers are relatively easy to locate and set up in any typical polling location. Voting software could be ported to various operating systems and various machines, providing flexibility to different precincts. While Internet voting is appealing for the reasons given above, several recent studies suggest that there are still some considerable security risks. Others include voters authentication technical problems, ballot secrecy, ballot integrity, reliable vote transport and storage, prevention of multiple voting, defense against attack on internet voting machines or election computer system.

1.1 Mobile Voting System (M-Voting)

M-voting is a further development of e-voting and telephone voting system. Owners of portable phones can subscribe to a service which enables them to participate in choosing political officials into government seats. M-voting is the use of mobile devices for citizens input to political decision-making. It is an m-government initiative with tremendous potential to enhance democratic participation and serves as an enabler for more convenient ways to involve citizens in political decision making. M-voting is attractive around the world as a way of encouraging participation, particularly among the young and in remote areas. It is also potentially far cheaper than other alternatives [8]. It is simpler to administer and obviates the need for polling places and warehousing of tabulation equipment. M-voting has been deployed in a number of regions in the developed nations these include: French Assembly chamber of commerce elections in 2005 with over 600,000 voters; City of Issy-les-Moulineaux (France): First French legally binding elections in 2003; Freie Hansestadt Bremen (Germany) Regional election (A Local government election); and Kista Stadsdelsnamnd (Sweden) [9].

M-voting is not the replacement of e-voting, rather it complements it. While mobile devices are excellent access devices, most of them, particularly mobile phones, are not suitable for the transmission of complex and voluminous information. Despite the emergence of more sophisticated handsets, mobile phones do not have the same amount of features and services as PC-based Internet applications [10]. However, there are more people who do not have access to PCs than there are people who do not have cell phones or other wireless devices [11]. M-voting also means that a citizen does not have to go and search for kiosks, or even get a connection to the house. People now carry a mobile access terminal with them wherever they go. Other benefits of using mobile devices for voting include: portability and mobility, flexibility, convenience, remote accessibility, ease of use and utility.

2.0 MOBILE TECHNOLOGY AND DEMOCRACY IN DEVELOPING NATIONS

Developing nations now recognize mobile technology to be a powerful tool for enhancing citizen engagement in public policy making. Many of these nations have begun to experiment with a range of ICTs (e-government and m-government) to enable greater citizen involvement in policy–making and the initial experiences illustrates the opportunities, dynamics and limits of these new tools [12]. Most governments are working to bridge the digital divide and recognize the need to ensure that all citizens whether online or not, continue to enjoy equal rights of participation in public sphere [13]. The initial lessons and experiences draw out the following points:

a. Mobile technology is an enabler for participatory democracy. The integration of these tools with the existing structures will make the most enhancing public participation in democratic processes [14].
b. The barrier to greater citizen participation in m-voting is cultural, organizational and constitutional but not technological. Overcoming these challenges will require

greater efforts to raise awareness and capacity both within government and among citizens.

The value of m-government comes from the capabilities of applications supporting mobility of the citizens, businesses and internal operations of the government. For example, supporting law enforcement agents who are on patrol is a distinctive advantage of m-government services over conventional government implementations.

2.1 Mobile Device Penetration

Mobile devices are now taking significant roles in our daily and business life with one third of the world population currently having access to mobile phone. This growth has been spectacular especially in European countries after the telecom industry de-regulation and adoption of Global system for mobile (GSM) communication [15]. Also, in Africa, particularly in Nigeria, with the introduction of mobile communication in 2001 mobile phones are no longer used only for voice communication but are a convenient way of connecting to the Internet. They are also used for transferring data, exchanging e-mails, decision making such as voting, and doing small business transactions. Next to increase in the adoption of mobile phones comes the growth rate of PDAs and Pagers. In 2001 the total sales of PDA was estimated to be well over 20 million worldwide [15].

A 2005 study by the Centre for Economic Policy Research and backed by the UK mobile phone giant Vodafone found higher rates of economic growth in developing countries with high mobile phone penetration. According to the study, a developing country which has an average of 10 more mobile phones per 100 populations between 1996 and 2003 would have enjoyed per capita GDP growth that was 0.59 percent higher than an otherwise identical country [16].

Over 85 percent of small businesses run by black individuals in South Africa rely solely on a mobile phone for telecommunications. The results of this study suggest that growth in the African telecom market will continue to pay off for African economies. In 2001, Africa became the first region where the number of mobile subscribers exceeded those using fixed lines.

Year	Subscribers
1998	2 million
2002	28 million
2003	51 million
2004	82 million

Table1: Number of cell phone users in Africa [17].

Mobile phones are particularly suited for developing countries such as Nigeria because, Internet access rate is low but mobile phone penetration is growing rapidly, particularly in urban areas where the number of mobile phones has surpassed the number of fixed/ wired phones. Total telephone lines as at 2003 in Nigeria, amounted to 853,100 [18], and mobile cellular phones in use as at December 2005 amounted to 19 million. Daily Trust newspaper of 16th March, 2006 records that about 20 million Nigerians have mobile phones [19]. Also research carried out by eShekels publication [20], estimates that mobile phone owners will grow to 28.8million by 2007. This number by far outweighs the total number of telephone line users. Furthermore, with the current deployment of 2.5G by operators in the country, the type of phones in present circulation may face a replacement demand.

The growth in mobile phone users has also been identified in many individual nations, including 49 middle income and 6 lower income countries. Among these countries are Burkina Faso, Chad, Honduras, Indonesia, Jordan, Mexico, Mongolia, Philippines, Saudi Arabia, and South Africa. According to a recent study, the population of global SMS users will grow to 1.36 billion in 2006 [10].

In Table 2 a comparison of the four main categories of mobile devices with their differing properties and penetration is shown.

Type of Device	Weight	Capability	Battery Life	Penetration
Mobile Phone	60-120 g	*	****	****
Handheld PC	90-200 g	**	***	**
Tablet PC	80 – 1200 g	***	**	*
Notebook	1500-400g	****	**	***

*= low, **= medium, ***= high, **** = very high

Table 2: Comparison of Mobile Devices [21]

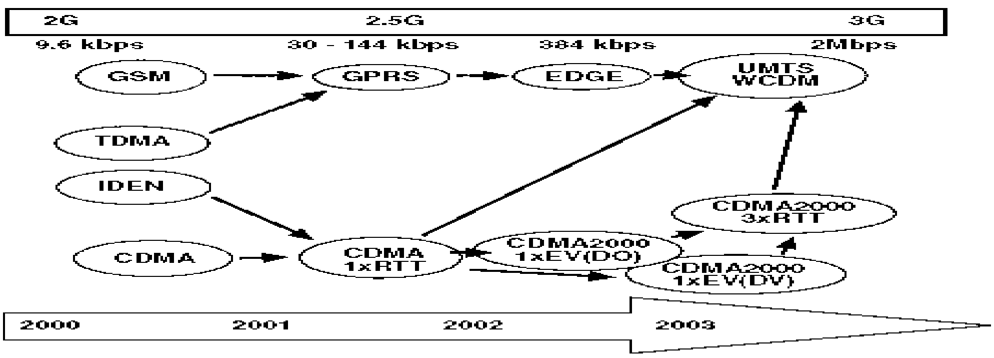
Figures released in March 2005 from the London Business School reported that Africa has seen faster growth in mobile telephone subscriptions than any other region of the world over the last five years. At the end of 2004 Africa's largest mobile phone firm, Vodacom, had 14.4 million users while MTN had 14 million subscribers. As Vodacom's chief tells Reuters, "Telecoms are Africa's big success story - perhaps the only one."

2.2 Major Concerns of Voting via Mobile Phones

1. The ability to convince the responsible political decision-makers of the advantages of m-voting. Therefore it is very important not to present the idea isolated from e-voting programs, but to put both of them in a close relation. M-voting and its tools have to be presented as a complement to e-voting and in addition to existing ways of participation.
2. The problem of persuading the users of mobile technology of the security of the system. This problem does not concern the whole democratic process, but above all the voting. Telecommunication companies and the federal authorities will have to work closely together and to work out an acceptable program [22].

2.3 Mobile Networking

One of the main components of m-voting system is wireless networking. Being wireless, gives the users opportunity to vote from anywhere the wireless network covers. There are different technologies for mobile networking and the bandwidths and speed rates of these technologies affects the types of applications in mobile networking. Figure 1 shows the different networks and bandwidth rates.



Fi

A ations and these includes; 1G, 2G, 2.5G and 3G. Figure1 and Table 3 shows examples of evolution of 2G networks to 2.5G and 3G networks and their data services.

Services	**2G**	**2.5G**	**3G**
E-mail	SMS	Text-based with small	Full attachment

		attachment	
Instant Messaging	SMS	Text-based	With Audio/Video clips
Web Browsing	Short text screens	100KB Web (text + image) page takes approx. 30 seconds to download	100KB Web page takes approx. 2 seconds to download
Streaming Audio/Video	No	Short clips	Yes
VoIP	No	Limited	Yes
File Transfer	No	500KB document takes approx. 2 minutes to download	500KB document takes approx. 10 seconds to download
Access to Corporate applications	Very limited	Text-based	Yes
Access to Corporate intranet, Databases	Very limited	Text-based	Yes
Location-based Services	No	Limited	Yes

Table 3: Examples of Data Services for 2G, 2.5G and 3G Networks [24]

2.4 Security Issues in M-Voting

User activities, weak username, and passwords, excessive permission, and users being deceived into revealing too much information can lead to security breaches in interactions between users and mobile devices.

Several mechanisms such as authentication, biometric authentication, smart card and so on can be used to alleviate these vulnerabilities. For the mobile client (mobile device), mechanism such as automatic logout, credentials re-entry, data destruction, database encryption and encryption of code-embedded usernames and passwords can be implemented to better protect mobile devices. Also code signing method should be used to ensure secure transfer of data to and from the server.

The wireless environment is less secured than the wired, and considering the importance of the voting process in a democracy, it is recommended that a third party (interswitch) be involved. This would forestall intrusion, denial of service attack and the 'gap-in-wap' syndrome

3.0 REQUIREMENTS FOR MOBILE VOTING SYSTEM

The development of any mobile application must address several types of requirements. Below are some lists of requirements for the mobile voting application development:

1. From the user's side, their mobile devices (mobile phones) are required which are already available. From the server side, servers and communication links will be required.
2. The capacity of the system will be made sufficient to deal with peak periods. Because network congestion could cause considerable frustration for people attempting to cast a vote [25].
3. The m-voting services should be such that the services are free of extra charges and also user-friendly as much as possible, so as to establish the citizens' confidence in the system and its applications.
4. It will be required that public law be put in place that this additional possibilities of politicking keep the same conditions as the traditional ways do. Legal restrictions have to be observed in areas such as: security and secrecy, integrity and

authenticity, and Data Protection Act. It is expected that adherence of the legal presuppositions will be of the biggest requirement of the system considering our polity.

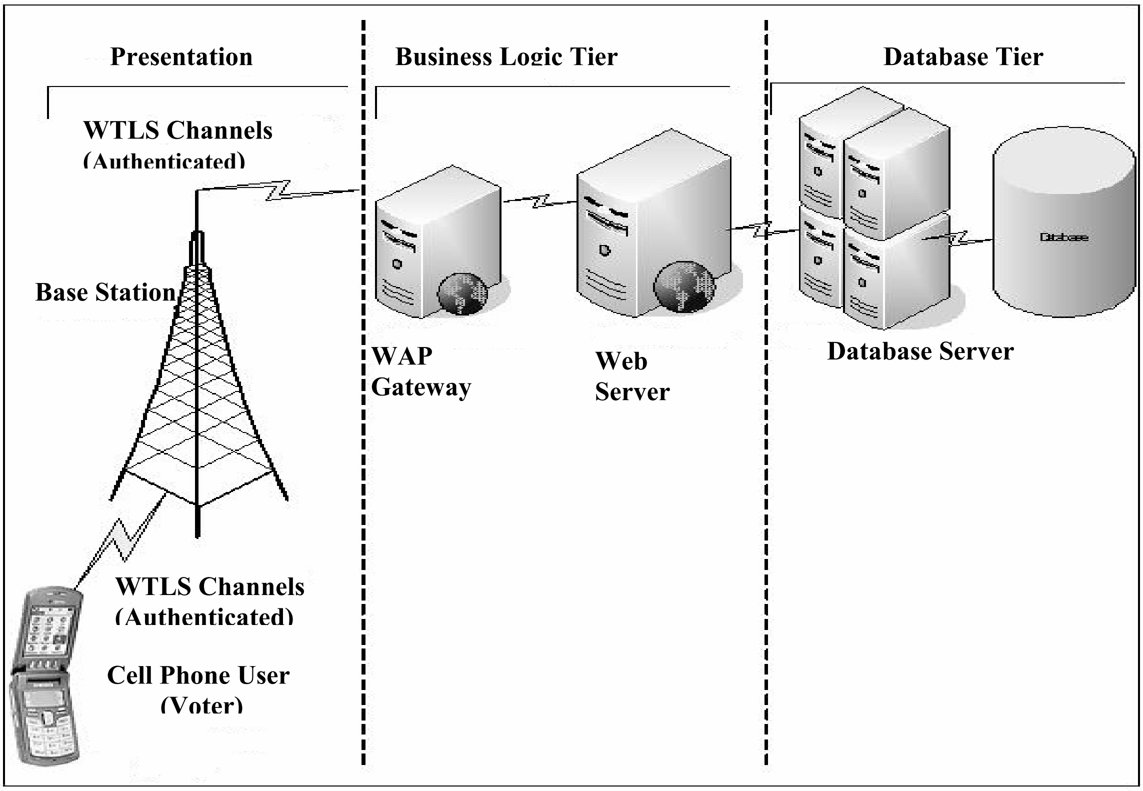
5. Election Requirements which include Eligibility and Authentication, Uniqueness, Accuracy, Verifiability and Auditability, Reliability will also be required.

4.0 THE M-VOTING SYSTEM ARCHITECTURE

The mobile voting application provides an example of mobilizing an existing web application so that it supports mobile devices such as the cellular phone. In this design, users are expected to have a constant network connection, and so we did not need to develop a fat client application for the mobile device. We merely have to create new WAP sites on the server that support a variety of mobile devices. It is expected that prior to the election, voters are physically registered at various designated registration centers where a voter's cards will be obtained. This is where the name, age, the thumb print, candidates ward, state of origin, local government area, sex, and other relevant voter information are captured. It is at this point that eligible voters are physically authenticated and registered. After registration the eligible voters obtain a voting card. The card will consist of a PIN code and a password which will be used to gain access to vote. Also, a code of five digits will be given to enable the voter gain access to the m-voting service. It is also expected that only WAP enabled mobile phones will be used. The m-voting system will be displayed on a mobile device (Cellular Phone).

4.1 System Architecture

The architecture of the mobile voting application is illustrated below. The mobile clients consist of the WAP browser (cellular phone). The mobile application consists of the following as depicted in figure 2, Presentation objects, Business objects, Data access objects and Databases. The mobile client has a zero application code layer on it. This means that it is a thin client. The server holds all the application code and it is organized as three-tier architecture.



This forms a simple architecture where the mobile client is assumed always to be connected to the server. Therefore there is no provision for storing application data on the mobile device. If the mobile device becomes disconnected, it will not be able to obtain up-to-date information until the connection is re-established.
The mobile clients are expected to cast their votes from different remote geographical regions. Each of these geographical locations will have its own server to ease congestion and easy administration during the voting periods. These servers collect all the data from that region and thereafter transfer it to the main voting server where the result will be collated.

5.0 CONCLUSION

When we weigh the requirements and the inherent roadblocks, m-voting appears to be relatively simple and a cost effective solution to open the citizen to ways of participation in elections. M-voting is easier to turn into action than e-voting projects; above all, the penetration of the mobile phone devices is much higher than that of the Internet, much easier for the users, and independent of location and time. The only risk that remains and which implementers should not run into is to consider m-voting as an isolated solution. M-voting is an ingenious voting initiative when considered in combination with or as a complement to other electronic voting systems. Furthermore, m-voting if adapted would enhance greatly the level of participation in elections as the elites who cannot stand the stress of voting in public domains can do so within the comfort of their homes. Similarly, the 3G licensing in the country will enhance the efficiency and reliability of data and video streaming through the WAP enabled mobile phone devices.

REFERENCES

1. Ayo C. K. (2006): E-DEMOCRACY: A Panacea For Enhanced Participatory Democracy, *The Information Technologist – An International Journal of Information and Communication Technology,* vol. 3, no. 1, pp. 43-53, ISSN: 1597-4316.
2. Ihenacho D. A, (2004) ‘Nigeria Dysfunctional INEC’, http://nigeriaworld.com/
3. Jones Douglas W. (2001) “Problems with computer based voting systems”, Http://www.cs.uiows.edu/~jones/voting/coloq/
4. Okop Umobong (2006) “Voting system in Nigeria”, A seminar paper presented at AEA, Blackpool England. Http://www.aea-lections.co.uk/annualseminars/downloads/nigeria-text.pdf
5. Bensor Kevin (2000) How e-voting will work, Http://www.howstuffworks.com/e-voting.htm
6. Rowell Erica (2002) “Man vs the machine”, Http://abcnews.go.com/sections/tech/dailynews/votertech001109.html
7. Heather Sterling (2002), “The impact of computer technology on voting” http://www.courses.cs.vt.edu/~cs3604/lib/Impact/Assignments/Voting.html
8. Steve Wallace (2005) “M-government Gets Serious”, http://www.thefeaturearhives.com/topic/technology/m-government-get-serious.html
9. Cybervote (2005) http://www.eucybervote.com
10. Emmanuel C. Lallana (2004) eGovernment for development, mGovernment definitions and models, Institute for development policy and management. Http://www.egov4dev.org/mgovden.htm

11. Michal Zalesak (2006) "M-government: more than a mobilized government", Web project Ltd., Zalesak
12. Macintosh, A, Davenport, E., Malina, A., and Whyte, A. (2002) Electronic government: design application and management, Umea University, Sweden, Idea group publishing.
13. Malina, A and Macintosh, A (2003) Bridging the digital divide: developments in Scotland.
14. Wilhelm, A (2000) Democracy in the digital age: Challenges to political life in cyberspace, New York and London, Routledge.
15. Sadeh, Norman (2002) "M-commerce: Technologies, services and business models", John Wiley and Sons, Inc. Canada and USA.
16. Rhett, Butler (2005) Cell phones may help save Africa, http://www.mongabay.com
17. ITU, (2002). Mobil Telefon _statistikleri, http://www.telkoder.org.tr/docs/mti-2.pdf
18. Wikipedia, the free encyclopedia.
19. Benjamin Aduli (2006) "Nigeria lose 0.5 million mobile phones annually over theft- survey", *Daily Trust* of March 16th 2006.
20. eShekel (2005) "Mobile Handset Report in Nigeria" 2nd edition.
21. Bullun, M., *et al* (2004) Egitimde Mobil Teeknolojiler, The Turkish online journal of technology, Vol 3, No. 2.
22. Heide Brucher and Petra Baumberger (2003) "Using mobile technology to support eDemocracy", Proceedings of the 36th Hawaii International Conference on System Sciences, Http://www.csdl.computer.org/comp/proceeding/hicss/2003/1874/05/1874504 4b.pdf
23. Ayo, C. K (2005) Lecture note on server side and mobile programming, Department of computer and information sciences, Covenant university, Ota (unpublished)
24. IBM Wireless e-business, (2001) Exploiting the full opportunity of 2.5G and 3G networks, http://www.c7.com/ss7/whitepapers/exploiting_3g.pdf
25. Coleman, *et al*, (2002) "Electoral reform society UK, Independent Commission on alternative voting method, elections in the 21st century; from paper ballot to e- voting", London.
26. Ekong, U. O (2006) Design and development of an M-Voting system, M. Sc Thesis, Covenant University, Ota Nigeria.

IMPLEMENTATION OF A LOW POWER SENSOR USING THE 1N4148 SIGNAL DIODE

F.E. IDACHABA AND V.E. EKONG

Department of Electrical and Information Engineering Covenant University Ota

+Department of Computer and Information Engineering Covenant University Ota

{ idachabafe@yahoo.com, victor_eshiet_ekong@yahoo.co.uk}

ABSTRACT

Wireless sensor networks are designed to be added to an existing infrastructure and to function remotely of it. They are characterized by limited radio and sensing range, the nodes are installed at a sufficient density to make it probable both that multidrop communication will be possible between any pair of nodes and that a significant phenomenon of the environment can be sensed. The networks have sensors for temperature, sound and light and they run on batteries and as such low power sensors are often desired for sensor networks. We propose a low temperature and low power wireless temperature sensor using 1N4148 signal diode that can ensure long term usage of at a significant power consumption and low cost.

Keywords: Wireless sensor network, transducers, diode

1.0 INTRODUCTION

A Wireless sensor network (WSN) consists of a number of small, low cost devices or nodes, each with facilities for sensing, computing and wireless communication ([1]. It is a special case of ad hoc networks where the nodes are physically arranged more or less randomly, but they can communicate over multiple wireless hops between their peers.

Wireless sensor network are designed to be added to an existing natural or built environment, and to function independently of it, without reliance on infrastructure. They provide distributed network and internet access to sensors, controls and processors that are embedded in equipment, facilities and the environment. These systems can provide monitoring and control capabilities for applications in transportation, manufacturing, environmental monitoring, safety and security. It combines micro sensor technology, low power and low cost wireless networking in a compact system. The lossless sensor networks led to the concept of smart environment. The smart environment relies on sensory data from the real world. Sensory data is collected from multiple sensors of different modulators in distributed locations [2].

2.0 WIRELESS SENSOR NETWORKS

Wireless sensors utilize small low cost embedded device and they do not need to communicate directly with the nearest base station. Peer to peer networking protocols are used to provide a mesh like inter connect that shuttle data between the embedded devices in a multi hop fashion. Figure 1 and 2 show the composition of wireless sensor networks.

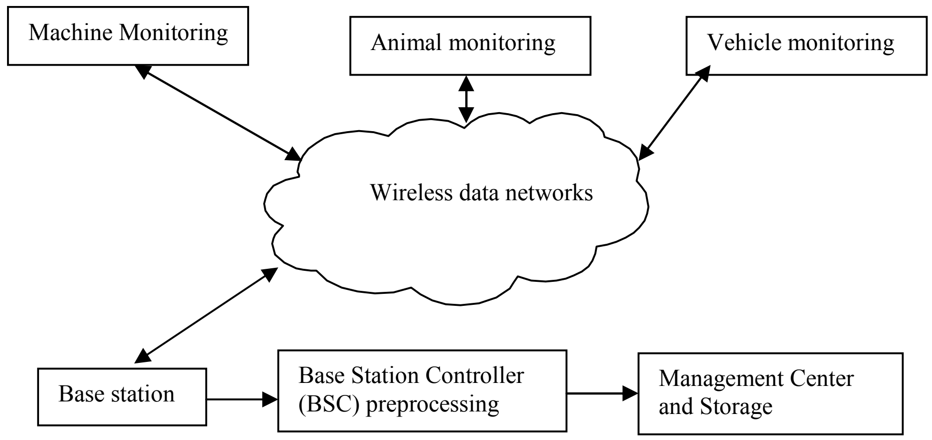


Figure 1: Data acquisition network

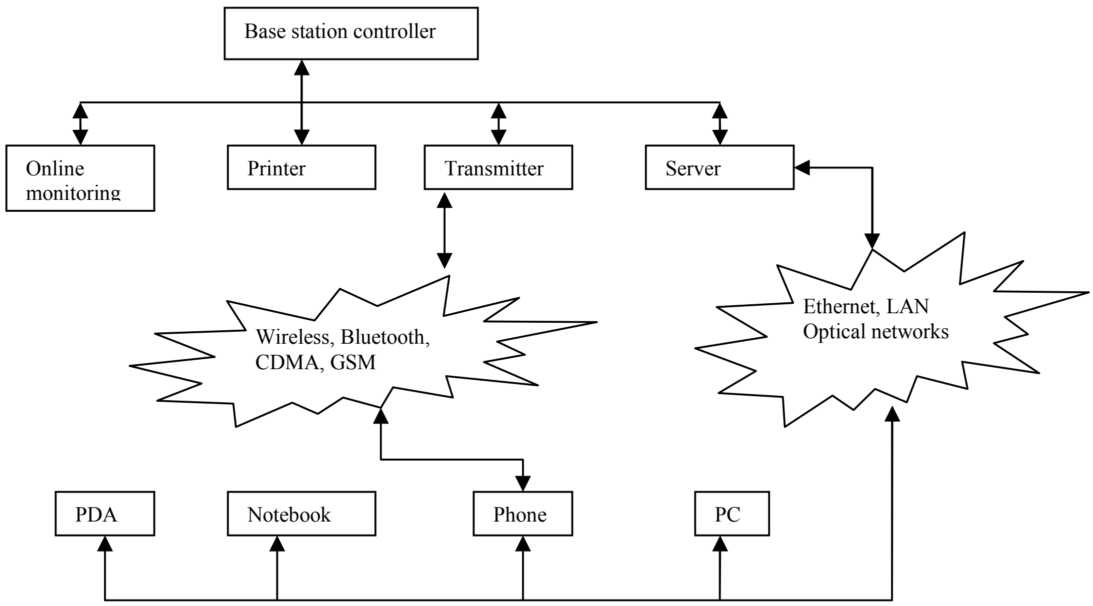


Figure 2: Data Distribution Network

The communication networks used with sensors is composed of nodes each with a computing power and transceiver capability over wired or wireless links.

Desirable functions for sensors include; ease of installation, self identification, self diagnosis, reliability. The IEEE approved in 1997 a smart sensor network standard known as IEEE 1451 to cope with compatibility of devices [3]

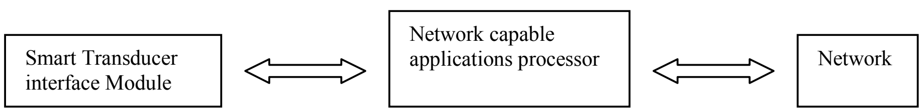


Figure 3: IEEE 1451 standard for smart sensors

This standard utilizes the principles of smart sensing. A smart sensor is a sensor which provides extra functions like signal controlling, processing and decision making/alarm functions. A general model of a smart sensor is shown in figure 4.

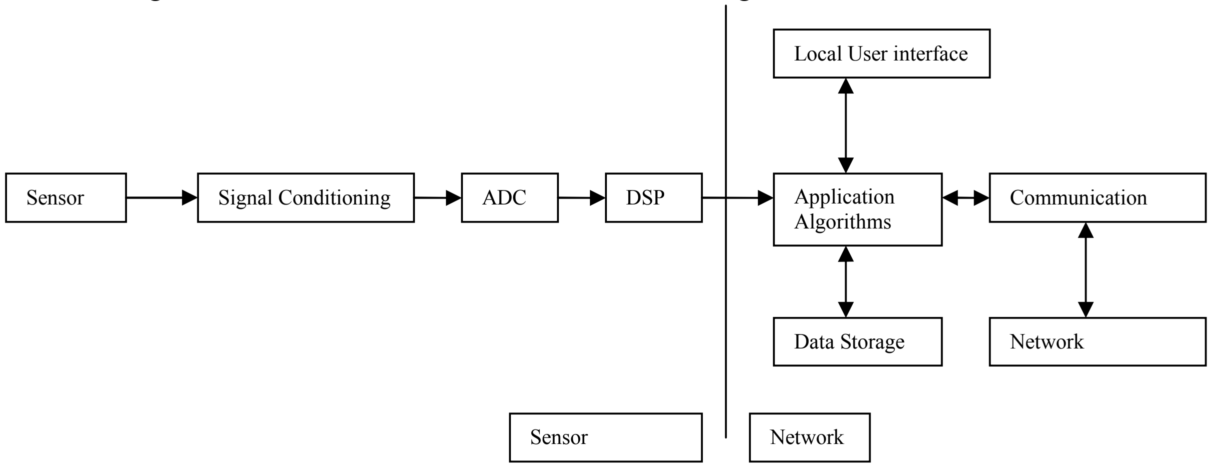


Figure 4 General model of a Smart sensor

We propose a model using an IEEE 4148 signal diode and transducers to design a low power wireless sensor network. A transducer is a device that converts energy from one form to another. The node (sensor) scattered on the environment, establishes a routing path and transmits data back in a collection point. If one node fails, a new topology will be selected and the overall network would continue to deliver data [4].

3.0 THE DESIGN ISSUES

Transmission Range

The wireless network will utilize omni-directional antennas as these antennas will allow nodes communicate effectively in all direction. Typical receivers are 85dBm and -110dBm. The transmission range for these device increases as the sensitivity and transmission power increases.

Topologies

There are three classic network topologies for sensors;

1. Point –to-point
2. Multi-drop and
3. Web network.

Point to point is the most reliable because there is only one point of topology and that is the host. (see figure 5). Each sensor node requires a separate twisted shielded pair connection. The cost of implementation is high and configuration management is difficult and almost all the information processing is done by the host [5].

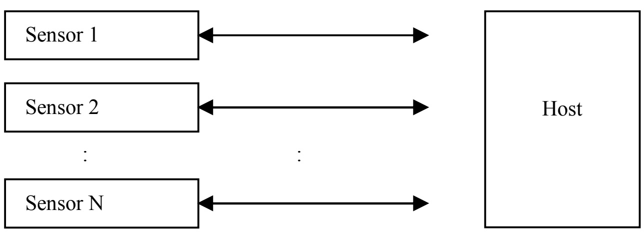


Figure 5: Point to point sensor network topology

In Multi-drop networks each sensor puts its information onto a common medium. Its use of a common medium introduced a potential single point of failure. Redundant connections are introduced to take care of this problem. The development of the Ethernet, Carrier Sense Multiple Access Schemes (CSMA) has enhanced the advancement of the multi-drop network.

The web network utilizes a small star network. The advantage of web connectivity for sensor network becomes clear as the level of intelligence in each sensor increases. Cooperating Sensors can form a temporary configuration to replace the host. Self- hosting network become self-configuring and then self aware. In a wireless web network, individual nodes have the potential of being connected (physically) with other nodes on the network. The configuration of the network is determined by the software running it. Routing is a major concern in web circulations. Repeaters are introduced in the network since all nodes cannot reach each other in a single hop [6], [7].

The Sensor

We use 1N4148 as a temperature sensor for the wireless network. The block diagram of this sensor is shown in Figure 6.

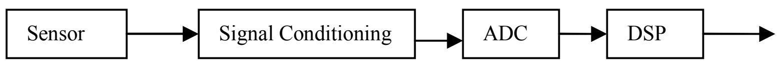


Figure 6: Block Diagram of the sensor

The IN4148 is considered for the wireless sensor application for a temperature ranges of 25^{o}C to 100^{o}C. The diode has the following advantages: -

1. Low cost
2. It is readily available
3. Rugged
4. Allows for simplicity of design

The diode current I_D is given as:

$$I_D = I_S\left(e^{\left(\frac{V_D}{nV_T}\right)} - 1\right) \qquad (1)$$

Where, I_s is the saturation current in the range of 10^{-15} to 10^{-13}A, V_T is the thermal voltage of approximately 0.026V and n is the ideality factor and ranges between 1 and 2.

When a constant current of about 1mA is passed through the diode, a forward voltage of approximately 600mV is developed in the junction. The precise value of the junction exhibits a negative coefficient of temperature of approximately -2mV/ ^{o}C.

The Sensor circuit [9]

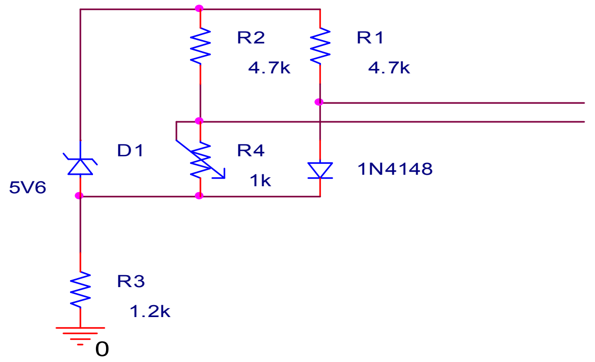


Figure 7: The sensor circuit

The current through the sensor diode (IN 4148) is given by

$$V = iR \tag{2}$$

Where V = Zener voltage 5.6V

$$i = \frac{5.6}{4.7k} = 1.19mA$$

The Zener diode is used to maintain a fixed and stabilized voltage across the ¼ bridge containing the sensor. The Zener diode ensures the development of a temperature dependent voltage across the diode.

Difference Amplifier

The sensor's output is amplified using an operational amplifier LM 741 configured as a differential amplifier. This circuit amplifies the difference between two input signals and rejects any signal common to both inputs.

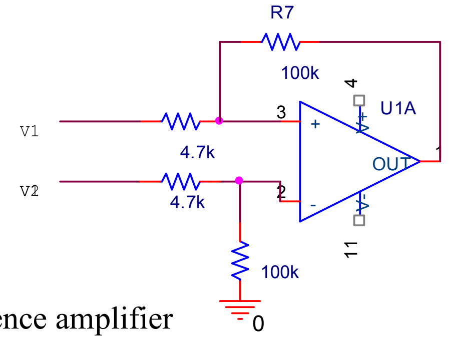


Figure 8: Difference amplifier

Applying the superposition theorem and the virtual earth concept;

R1 = R3 = 4.7k, R2 = R4= 100k

$$V_{01} = \frac{-R_2}{R_1} V_1 \tag{3}$$

$$V_{02} = \left[1 + \frac{R_2}{R_1}\right]\left[\frac{R_4/R_3}{1 + R_4/R_3}\right] V_2 - \left[\frac{R_2}{R_1}\right] V_1 \tag{4}$$

An important property of the differential amplifier is that its output voltage is zero when both inputs are the same (common mode rejection).

This forms the equation above, with $V_1 = -V_2$

$$\frac{R_4}{R_3} = \frac{R_2}{R_1} \tag{5}$$

$$V_0 = \frac{R_2}{R_1}(V_2 - V_1) \tag{6}$$

The differential gain is then given as:

$$\frac{R_2}{R_1} = Ad \tag{7}$$

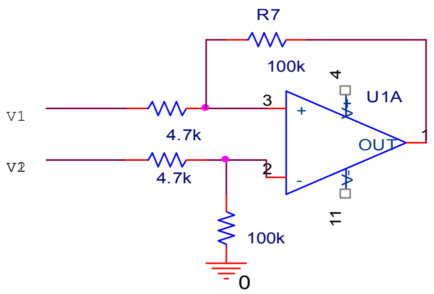
The common mode gain is defined by A (cm) = $\frac{V_o}{V_{cm}}$

The common mode rejection ratio is used as a quality factor for differential amplifier. Good differential amplifiers have a CMRR in the range of 80dB to 100dB.

$$CMRR = \frac{Ad}{A_{CM}} \tag{8}$$

$$CMRR(dB) = 20\log\frac{Ad}{A_{CM}} \tag{9}$$

The differential amplifier circuit is shown below;



$$Ad = \frac{100k}{4.7k} = 21 \tag{10}$$

The sensor circuit combines with the differential amplifier circuit to yield the transducer circuit as shown in figure 9.

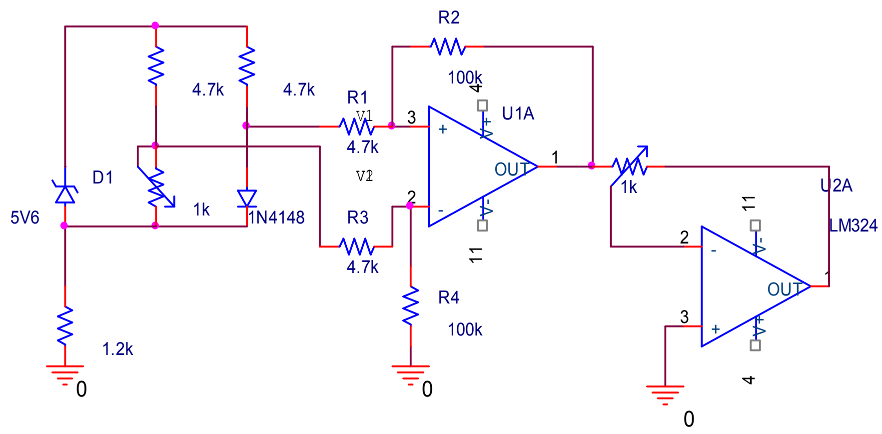


Figure 9: Transducer circuit

From equation (4);

$$V_0 = \left[1+\frac{R_2}{R_1}\right]\left[\frac{R_4/R_3}{1+R_4/R_3}\right]V_2 - \left[\frac{R_2}{R_1}\right]V_1 \qquad (11)$$

$$With \frac{R_4}{R_3} = \frac{R_2}{R_1} = \frac{100k}{4.7k} = 21$$

$$V_0 = (1+21)(\frac{21}{1+21})V_2 - 21V_1$$

$$= 22(0.95)V_2 - 21V_1 = 21V_2 - 21V_2$$

$$= 21(V_2) - 21(V_1) = (V_2 - V_1)(21-21).$$

$$V_0 = 0$$

Differential input voltage is given by;

$$Vd = V_2 - V_1 \qquad (12)$$

$$V_{CM} = \frac{V_1 + V_2}{2} \qquad (13)$$

$$V_1 = V_{CM} - \frac{Vd}{2} \qquad (14)$$

$$V_2 = V_{CM} + \frac{Vd}{2} \qquad (15)$$

Substituting equation (14) and (15) into (11);

$$V_0 = 21(V_{CM} + Vd/2) - 21(V_{CM} + Vd/2) \qquad (16)$$

$$V_0 = 21(Vd) + 0V_{CM}$$

This output voltage is given to be of the form

$$V_0 = AdVd + A_{CM}V_{CM} \qquad (17)$$

From equation (16) and (17)

$A_d = 21$ and $A_{cm} = 0$

From the above it is observed that the CMRR is very high.

4.0 EXPERIMENTAL TEST-BED

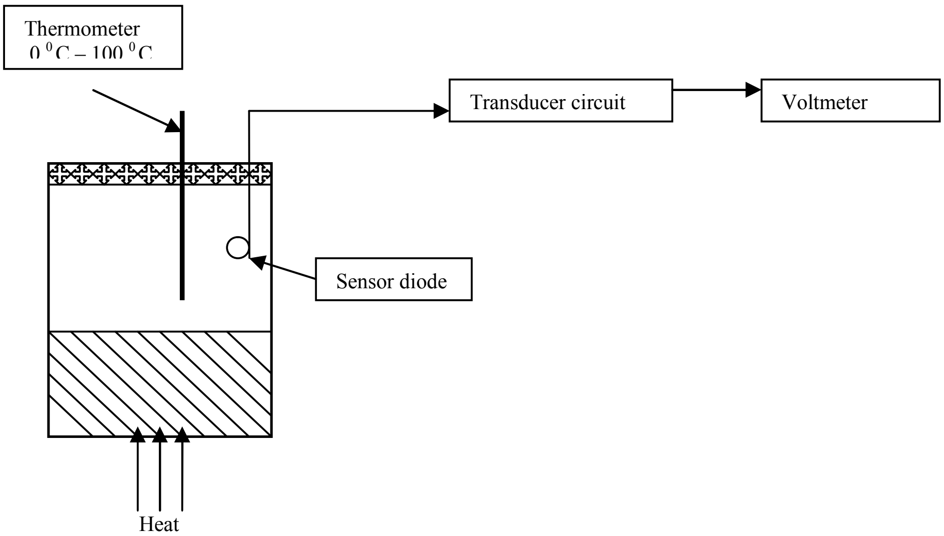


Figure 10: Experimental test set up [8].

The proposed set-up above is used to determine the temperature characteristics of the sensor. The water in the container was heated from room temp.(27°C) to boiling point (100°C) with a transducer voltage reading taken for every 5°C rise in temperature. The water is allowed to cool down and the cycle repeated. Ten different readings were taken and the average of the readings is presented in the table below.

Table 1 Temperature reading of the sensor diode

Temperature(°C)	Transducer output voltage(V)
27	3.96
30	3.92
35	3.83
40	3.69
45	3.55
50	3.40
55	3.24
60	3.05
65	2.86
70	2.67

75	2.47
80	2.27
85	2.05
90	1.76
95	1.49
100	1.33

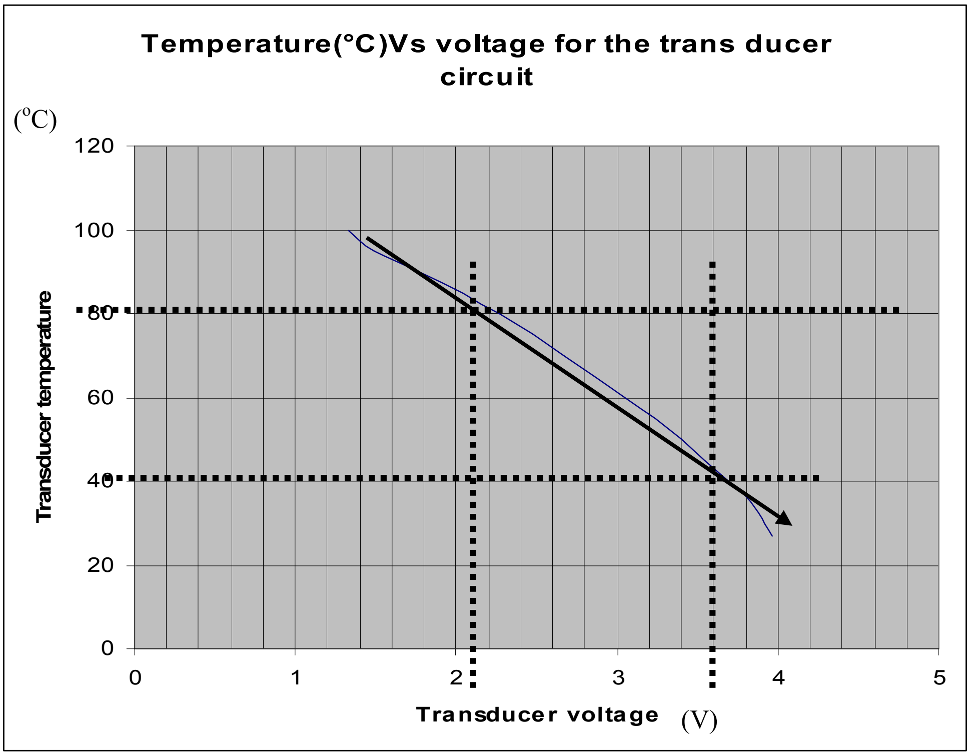


Figure 11: Graph of results for temperature and voltage readings taken

The slope of the transducer circuit is given to be:

$$\text{Slope} = \frac{2.67 - 2.27}{70 - 80} = \frac{0.4}{-10} = -0.04\ \text{V}/^{0}\text{C} = -40\text{mV}/^{0}\text{C}$$

The readings show the negative temperature coefficient of the diode sensor and the influence of the differential amplifier. An inverter circuit can be added to give the transducer output a positive temperature coefficient and bring the voltage readings to a range suitable for ADC chips. The resultant transducer circuit and the readings are shown below:

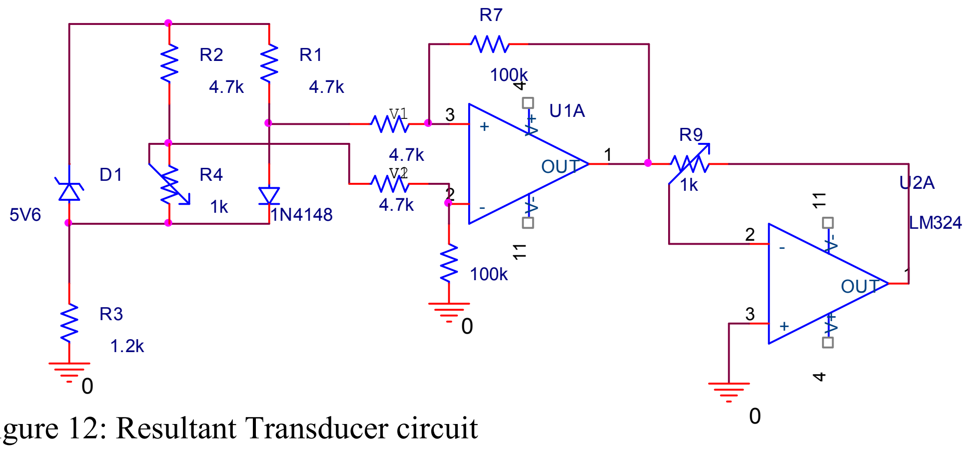


Figure 12: Resultant Transducer circuit

Table 2: Resultant system output results

Temperature ^{o}C	Transducer output	Inverter output voltage
27	3.96	1.31
30	3.92	1.32
35	3.83	1.33
40	3.69	1.35
45	3.55	1.39
50	3.40	1.43
55	3.24	1.48
60	3.05	1.53
65	2.86	1.59
70	2.67	1.64
75	2.47	1.70
80	2.27	1.76
85	2.05	1.82
90	1.76	1.90
95	1.49	1.99
100	1.33	2.04

5.0 CONCLUSION

From our design and the test bed readings obtained, it is observed that the silicon diode can be used a as sensor within the range of 27°C to 100°C conveniently. The overall sensitivity and temperature coefficient can be altered by the use of a suitable circuit. The resultant circuit can be realized using minimal discrete component or completely fabricated in integrated circuit (IC) chips. The final sensitivity of the overall circuit using the readings taken from the inverter output shows the sensitivity as given below.

$$\text{Sensitivity} = \frac{1.53 - 1.43}{60 - 50} = 0.01 = 10\text{mV}/^{o}\text{C}$$

This shows an increase from the -2mV/^{o}C of the diode and a conversion of the temperature coefficient from negative to positive which simplifies further signal processing.

REFERENCES

[1] George C., Jean D. and Tim, D (2005) *Distributed systems: Concepts and design*, 4th ed., Addison Wesley, England.

[2] Lewis, F. L. (2004) *Wireless Sensor Networks: Smart Environments Technologies, Protocol and Application*. D. J. Cook and S. K. Das (*Ed*) John Wiley New York.

[3]. Culler, D., Estrin D. and Srivastava M. (2004) Overview of sensor network, *IEEE Computer*, Vol. 37, no. 8, August, pp. 41-49.

[4] Romer, K. Oliver K., and Friedmann M. (2003) Middleware challenges for wireless sensor networks, *Mobile Computing and Communications Review*, Vol. 6 No. 2.

[5] Hill, J. L (2003) System Architecture for Wireless Sensor Networks, Ph.D Thesis University of California, Berkeley Spring.

[6] Ekong, V. E (2002) Wireless communication network design issues, M.Sc Thesis, University of Benin, Benin city, Nigeria.

[7] Culler, D., Hill, J., Buonadonna, P. *et al* (2001) A network-centric approach to embedded software for tiny devices, *Proc. of the first International workshop on embedded software*, Tahoe city, CA USA, pp. 114- 130.

[8] Idachaba, F. E (2002) Design of a microcontroller based process monitoring and control unit using the 8751 microcontroller, M.Eng Thesis University of Benin, Benin City, Nigeria.

[9] Donald A. Neaman (1996) Electronic circuit analysis and design, McGraw-Hill, USA.

IMPLEMENTING OPERATIONS SUPPORT SYSTEMS IN E-HEALTH BASED SYSTEMS

N.A. IKHU-OMOREGBE, S.A. EHIKIOYA, C.K. AYO

Department of Computer and Information Sciences,
Covenant University, Ota, Nigeria

ABSTRACT

Information and communication technologies have been introduced in different dimensions of the health care. e-Health is the use of advanced communications technologies such as the Internet, portable, wireless and other sophisticated devices to support health care delivery and education. It has the potentials of improving the efficiency of health care delivery globally.

With the increasing demand for information at the point of care, health care providers could explore the advances provided by mobile technologies and the increasing capabilities, compactness and pervasiveness of computing devices to adopt operations supports systems (OSS) in e-Health based systems in order to provide efficient services and enhance their performances.

In this paper, we present, the development and implementation of operations supports in e-Health based systems. The system promises to deliver greater productivity for health care practitioners.

Keywords: Architecture, e-Health, globally, portable, productivity and wireless devices

1.0 INTRODUCTION

Over the last decade, the need to develop and organize new ways of providing efficient health care services has resulted in a dramatic increase in the use of Information and Communications Technology (ICT) based solutions in health care delivery, generally known as e-Health.

This entails a fundamental redesign of health care processes based on the use and integration of electronic communication at all levels. A benefit of e-Health is that patients in one country can seek treatment and other services from other countries with their patient records moving with them electronically. A patient with his medical information moving electronically via wireless technology is empowered to play active role in the decision making process during his treatment. Treatment processes in care centres are facilitated with the introduction of operations support systems (OSS).

e-Health based OSS are essentially designed to automate manual health care processes, making the operations of health care practitioners and the network infrastructure more error-free and efficient. OSS solutions for e-Health takes advantage of state-of-the-art information technology to address health care enterprise-wide needs and requirements to reduce costs, provide reliable, flexible, mobile, timely, secure health care delivery to patients by health care providers, centres or practitioners

Research efforts and the use of communications technologies (fixed and mobile) to extend the reach, range and manoeuvrability of health care application and contents are covered in [1, 2, 3, 4]. It is not uncommon in recent times, to have a number of personalized applications for prescription, research, reference, patient education, and other applications for accessing daily schedules being bundled with hand-held devices to provide supports for physicians [5, 6, 7]. However, research on applications based on wireless LAN to improve health care services in the hospitals are becoming more popular [1, 3, 8] because of the increased demand for real time access to medical information by medical and support personnel, the maturation of 802.11 standards, and the proliferation of 802.11-enabled devices. Simple wireless devices such as PDAs from literatures [9, 10, 11] are reported to be more flexible and portable for use by physicians than some more computational desktop computers.

The use of mobile technology devices such as PDAs, cell phones, laptops, etc for health care (mobile health care) [12] delivery promises a revolution to modern health care as physicians stand to make more current and accurate prescribing decisions, thereby reducing harmful drug interactions. In addition to enhancing quality of patient care, e-Health technology, through the elimination of redundant paperwork, also facilitates more efficient and effective delivery of patient care.

The remaining part of this paper is discussed as follows; In Section 2, we present a design of the system, in Section 3, we present the implementation technologies and a discussion of users' interactions with the system in Section 4. We conclude in section 5.

2.0 SYSTEMS DESIGN

Prior to our design and implementation, we accomplished our requirements engineering processes by interacting with the various medical practitioners at the Covenant University Health Centre to gather the requirements for the system. We were able to identify some supports services which are represented architecturally in the following sections:

2.1 Operations Support Services in e-Health Systems

The supports services and their main features are discussed as follows;

- Patient Medical Record Support Service (PMRSS)

  This is the core support service in e-Health based systems [13]. It is one of the components in the application layer of the software architecture and interacts closely with the laboratory test support service(s), and the pharmacy and billing support service(s). It communicates with other support services in the application layer to capture the required data which are committed to the data layer for more permanent storage. It solves many of the logistic problems of archiving, retrieving, tracing and finding out paper based medical records. It provides functionality that maximizes the usage of the patients' medical information by serving as central sources of information for communications between health care providers, covering the patients' history, observations, diagnosis and therapeutic conclusions.

  With PMRSS in place, reports such as laboratory test, medication, doctor-patients appointment, billing information, patient and staff data are easily generated.

- Laboratory Test Support Service (LTSS)

This support system aims at improving patient safety and efficiency in care delivery. A doctor who wants to confirm the presence or absence of diseases may request that a patient undergoes a laboratory test. After the test is performed, the result is given to the patient or captured and committed directly to the system via a mobile device for a later use by the referring physician.

- Pharmacy and Billing Support Service (PBSS)
  This module is patient-centred. Its functionalities include checks for drugs interactions, billing and allergic drug alert. The module helps to minimize medication errors and adverse drug effect on patients by allowing physicians send prescription information directly to the pharmacist after reviewing patient's allergic medication list real-time via handheld devices. It incorporates a knowledge driven risk management support to enhance clinical decision support from the point of care. The system automatically generates the bills and forwards prescription information to the central server for review and dispersal by the pharmacist in due course.

- Documentation Support Service (DSS)
  These support services enable authorised users especially physician to record complaints, diagnoses, allergies, medications, referral information and other data electronically. As patient visits to a care centre, the nurse or medical record personnel calls up his details from the e-Health system in readiness to provide appropriate services. Where his record does not exist, a new record is created with the patient demographic information, vital statistics etc. The patient registration process and other encounters or interactions with medical personnel are equally captured with standard interfaces.

- Data Access and Update Support Service (DAUSS)
  These support services provide an interface for authorised users to easily retrieve and modify information. Each patient record is assigned a unique identity number. Patient identification and verification is required to maintain data integrity of the system. This is achieved by keeping a unique patient's t Identity across the entire patient's records. At the point of care, the physician is able to search for a patient's record in relation to a service earlier received by providing the identity number. The system equally provides support for the modification or update of records where necessary.

- Security and Authentication Support Service (SASS)
  In a hospital setting, several users access the patient's records to improve services to patients. These users include, doctors, nurses, laboratory personnel, pharmacists amongst others. A common way to give access right to hospital personnel is according to their functional and structural role which relates to their profession or specialty [13]. Based on the level of authorization, each personnel is allowed to view and modify only those information he is authorized to have access to.

- View and Report Support Service (VRSS)
  Electronic view capability is vital for e-Health bases systems. At various points in a patient care, it may be necessary to review or visit a patient's medical information. To view a patient's record, an authorised user after been authenticated by the system, is allowed to search for patient's record by providing the patient's identity. The reason for a view may include a review of doctor's diagnosis, past medications, laboratory result or test, bill information. An update is allowed where wrong information was earlier captured or entered.

The process of treating a patient requires some mobility of the patient or the medical personnel while the patient record moves electronically. The cooperation between the different personnel is captured in the collaboration diagram in Figure 2.1.

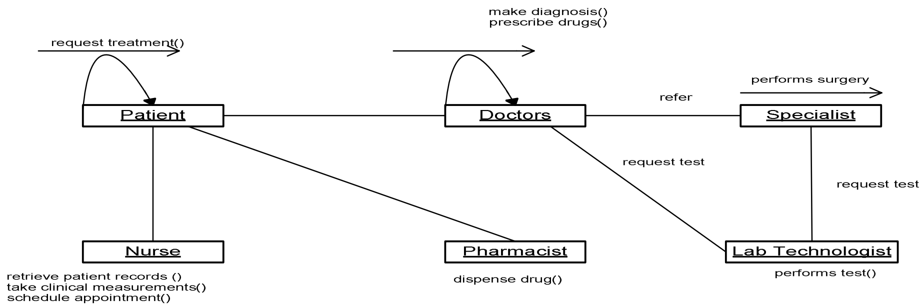


Figure 1: Collaboration Diagram

2.2 Architectures

Health care delivery in a hospital environment is inherently mobile combined with the tremendous need to have timely and accurate access to patients' information at all times. Thus the architecture proposed must allow easy capture, search, retrieval, and update of patient information from the point of care within the hospital setting. We used a client-server multi-tier based architecture for the hardware (physical) and a 3-tier client / server architecture for the logical software (logical) implementation and deployment. The two architectures are depicted in Figure 2 and Figure .3 respectively.

2.2.1 The Hardware Architecture

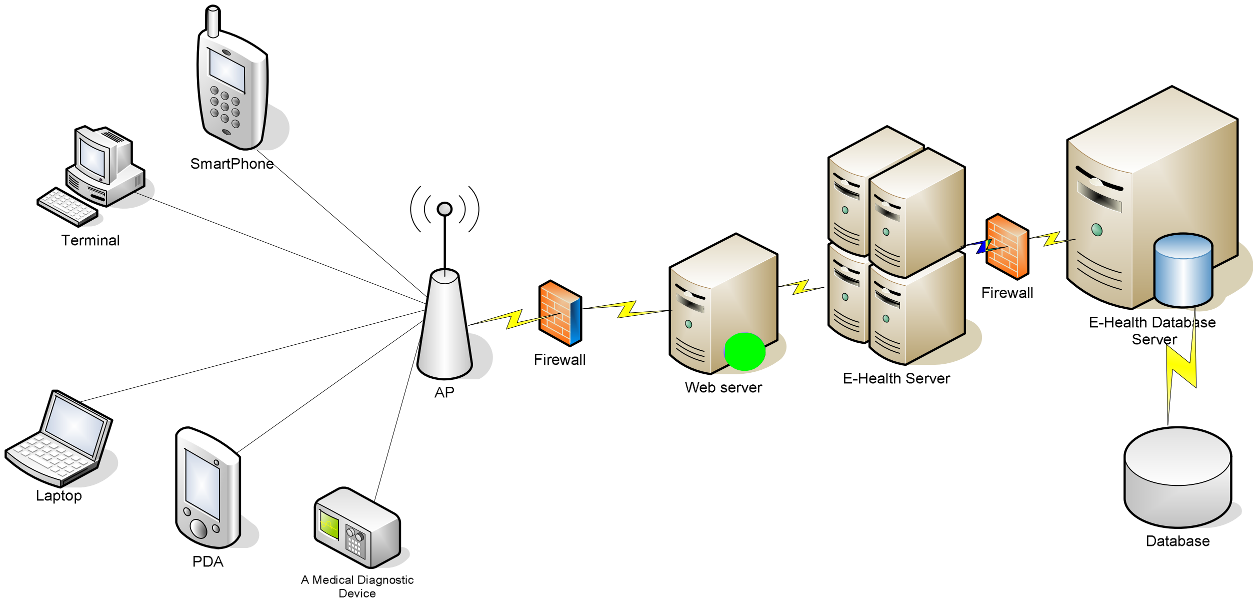


Figure 2 : Hardware Architecture for e-Health Based System

This architecture consists of a complete range of robust performance client and server platforms with integrated enterprise application and data extendable to care providers at the point of care.

The client systems include PDA, cell phone and smart phone and other handheld medical devices that combine real time access to enterprise systems and data with processing power for local analysis. The client devices have features to connect to enterprise resources and application over wired LAN, 802.11 based wirelesses LAN or high speed wireless wide area network with viewing surface ideal for recording, searching, analysing and reviewing patients' information.

The servers are used to maintain connectivity to enterprise resources for the mobile health care providers that include the doctors, nurses, laboratory technologists and other support staff. Load on the application server tier is balanced by using multiple application servers.

Firewalls are set up to filter all network traffic moving in and out of the e-health system. The hardware architecture is highly secure and utilizes multiple layers of firewall protection to create several region of trust. The robust servers provides, real-time access to point of care database originating from systems across the enterprise system within the mobile environment to facilitate timely and accurate care delivery and practice management.

2.2.2 The Software Architecture

Figure 3 gives an overall logical (software) view of the architecture of the e-Health system. The architecture shows the locations of the each of the supports service in the system. It is a 3-tiered client-server which consists of the client interface, middleware and database repository. The database is separated from the client through the middleware. The middleware concept helps to solve scalability, load balancing, transactional processing and interoperability issues by providing a means that allows different hardware and software from different manufactures to share common Patient medical records.

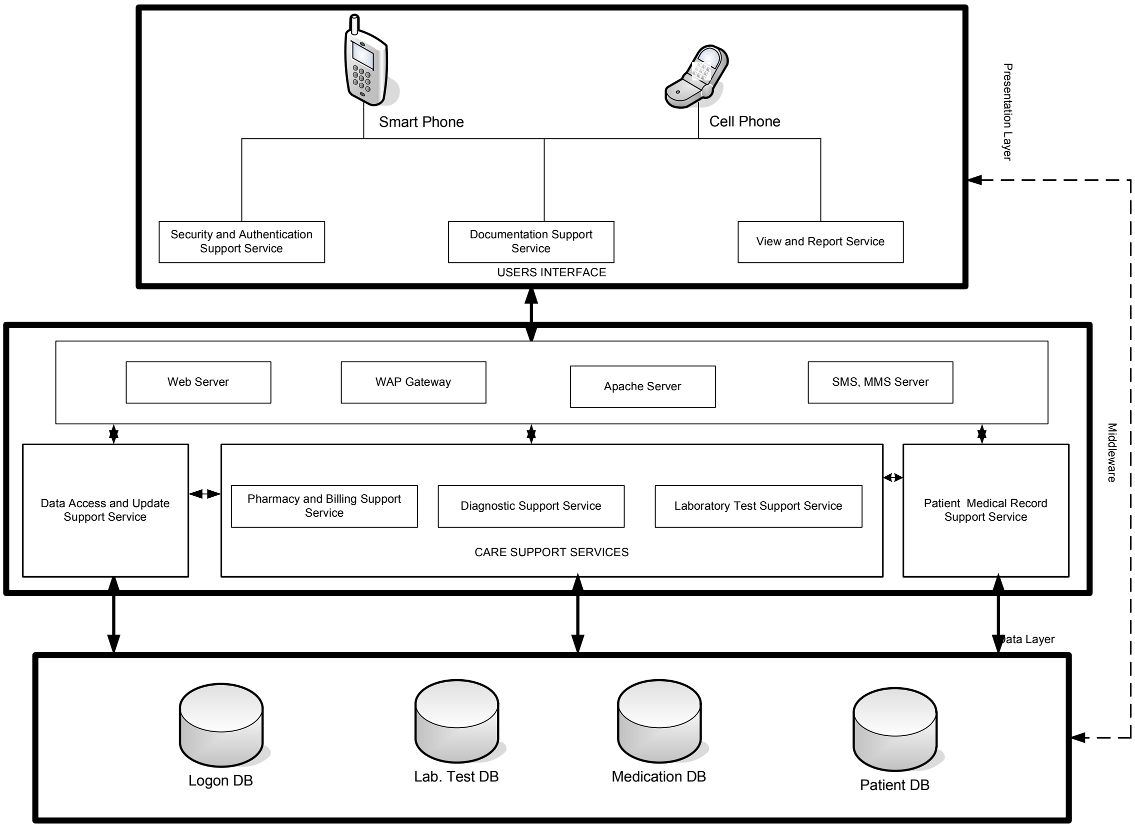


Figure 3: Software Architecture for the e-Health System Implementation

The mobile clients have a zero application code layer on it, that is, it is a thin client. The client has as its components the Security and Authentication Support Service (SASS), Documentation Support Service (DSS) and the View and Report Service (VRS). These support services do not store or process any form of data. They only provide an interface for middle layer and the data layer.

3.0 IMPLEMENTATION TECHNOLOGIES

The prototype application was implemented and tested on a client-server architecture separated by a mobile network. The server application developed in Java was deployed on an Apache Server running on a Windows Operating System. Java was used for the application development to ensure it portable across various platforms. The client application accessed the server application via a windows CE micro browser on a PDA.

The design of WML decks required special considerations due to the resource constraint associated mobile devices. The application starts with WML cards that introduce the user to the support services in the application. After a deck has been downloaded a copy of it is cached in the mobile device for some time depending on the system's configuration. Java Server Pages were used in addition to WML to add dynamic functionalities to the static WML pages by providing access to an MS Access database through ODBC-JDBC Bridge.

Openwave V7 Simulator provided a cost effective platform for testing the application at the development stage as it is free on the Internet. At the implementation and testing stages the application was deployed on an **O_2 Xda Mini S PDA running** Windows Mobile 5.0 Operating System. The WAP platform facilitated the delivery of contents from the Server to the mobile device via a WAP gateway. The WAP content compiled at the gateway are sent to the device as binary content.

4.0 DISCUSSION

The following screen shots show the interactions of the users with the application.

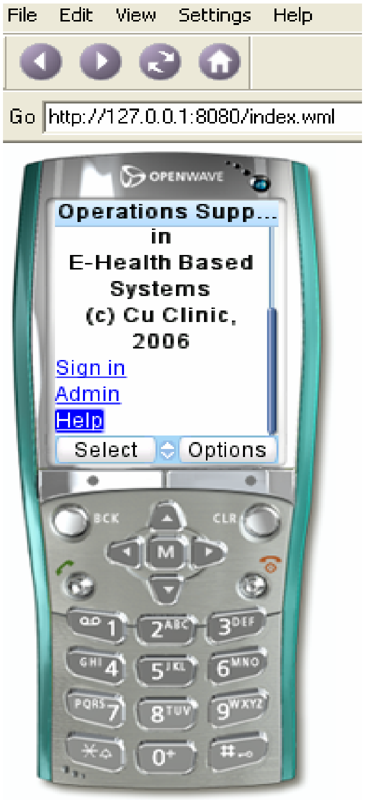


Figure 4: Welcome Page

The application is launched with a welcome message and the hyperlinks to some of the functionalities within the application. To facilitate the efficient and effective mobile health care, it is essential to provide the functionalities which directly support the user in their preferred way of performing their task. For example, the interface provided in the doctor's module allows him to capture each of the variables in a usable pattern. Figure 5 shows a medical doctor's interaction with the system.

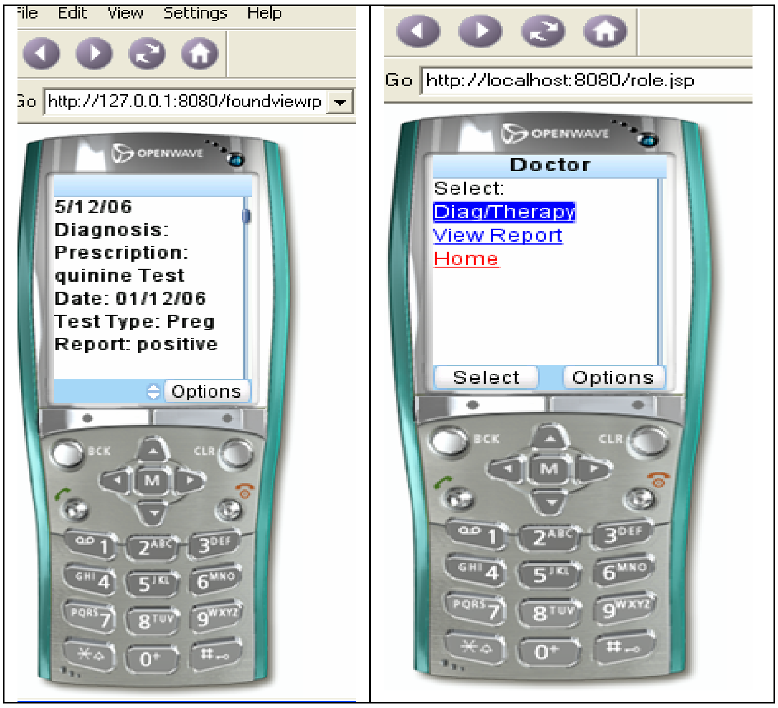


Figure 5: Doctor's Interaction

CONCLUSION

In this paper, we are able to demonstrate the implementation of an application for mobile devices such as mobile phones and PDA in enhancing the effectiveness of health care professionals in the delivery of services to patients. The various supports services were adequately depicted in robust architectures for easy implementation.

The deployment and adoption of this application will improve medical services which are presently cumbersome with so much of paper work in most health care centres in the country. Furthermore, it will help many countries to meet their national health care policies and the millennium development goal 3, goal 4, and goal 5 which bother on health care.

REFERENCES

[1] S. Voskarides, C.S. Pattichis, R. Istepanian, E. Kyriacou, M.S. Pattichis, C.N. Schizas., "Mobile health systems: A brief overview" Available at: www.medinfo.cs.ucy.ac.cy / documents / spie02.pdf

[2] "Documenting Component and Connector Views with UML 2.0", *Technical Report*, CMU / SEI-2004-TR-008, ESC-TR-2004-008

[3] James E. Cbral Jr. and Youngmin Kin, "Multimedia Systems for Telemedicine and their Communication Requirements", *IEEE Communication Magazine*, 1996

[4] Asfanyar Qureshi, Ali Soeb and John Guttang, "Building a High-Quality Mobile Telemedicine System Using Network Striping over Dissimilar Wireless Wide Area Network"

[5] The Journal of Medical Informatics, Available at: http: //www.rnpalm.com /

[6] "PDA Cortex store", Available at: www.pacortex.pdaorder.com

[7] Ravi Raman, Ramana Reddy, Sumitra Reddy, "Collaboration framework for Nomadic Health care Provide", *Proceedings of the Eleventh International*

Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE02), IEEE, 2002

[8] Y.-C. Chen, H.-C. Chiu, and L. Liu (Taiwan), "The Consultation Specialist System of PDA in Mobile based e-Hospital", *Biomedical Engineering* , 2005, Available at: www.actapress.com

[9] B. W. Trevor Rohm and C.E. Tapie Rohm, "Evolving medical informatics: from diagnosis to prognosis", *Int Journal Electronic Health care Vol. 1. No 1, 2004*

[10] Wickramasinghe, N and Misra S. K., "Wireless Trust Model For Health Care", *International Journal Electronic Health care* , 2004 Vol. 1. No1. pp 60-77

[11] Darren Mundy and David Chadwick, "Electronic Transmission Of Prescriptions: Towards Realising The Dream", *International Journal, Electronic Health care,* Vol 1, pp. 112-125

[12] Olubukola D., A. and Ayeni J.O.A., "Human-Computer Interactions Factors: Usability and User experience in Software Designs", Information Technology Capacity Building: The Future of Nigeria's Economic Growth (ICABUILD2005)*Conference Proceedings, Volume 16, 8th International Conference, Nigeria Computer Society, 14-17th June 2005, Port-Harcourt, Nigeria, pp 79.*

[13] Ragnar Nordberg, "EHR in the perspective of security, Integrity and Ethics", Medical and Care Compunetics 3*, L. Bos et al (Eds.), IOS Press, 2006*

BROWSING THROUGH THE FOURTH GENERATION MOBILE NETWORKS

O. SHOEWU, O. BADEJO AND O.T. BARUWA[1] **AND D.W.S. ALAUSA**[2]

1. Department of Electronic and Computer engineering
Lagos state university, epe campus, epe

2. Department of Computer Engineering
Federal Polytechnic, Ilaro, Ogun State

ABSTRACT

This paper presents the evolution of mobile networks pointing at trends and tendencies. By 2008, the first signs of fourth generation mobile systems will appear. An account of the user's requirements as of now and how it will look like in the future shall be discussed. The goal of this paper is to confirm the necessity of the fourth generation systems and devices.

Keywords: Mobile Networks, Digital Technology, Data Transmission, Modulation, and Frequency Bands, Fourth Generation.

INTRODUCTION

The introduction of the third generation mobile networks (Zeng, et al. 2000) has just surfaced, and it holds out the possibilities of what have never been seen before. All mobile generation created something new and the fourth generation will not be an exception. What can be the expectation now, when we have not even reached the borders of the capabilities of the first generation in Nigeria and the second-generation networks worldwide, and the notion of 3G network has just become generally known? To answer the questions highlighted there will be a need to browse through the phases of the improvement of mobile systems in the ensueing discussions.

COMPARISON OF AD HOC NETWORKS

The first generation mobile networks were developed in the early eighties; they were based on direct analogue modulation and adapted cellular structure. Their main desert was the realization of the cellular concept by reusing frequencies. Circuit switched communication was an idea prevalent in them. The idea of mobility and the maintenance of nimbleness appeared in this period among the users while cooperation of independent systems was not accomplished. However, the roaming function was also missing.

The second-generation mobile networks could owe their success and progress to the intensive development of microelectronics and the digital technology. As the result of this development, the weight of the terminals (handset) considerably decreased (from about 1killogram to between 10-20gram). Due to their more user-friendly structure, the appearance of the SIM card and their services based on circuit switched digital communication, millions of user opted for this successful system and chose a GSM terminal (Shoewu, 2003).

In the frame of circuit switched network structure besides the transmission of voice another possibility arose to low- (9.6kbps, 144.4kbps) and high-speed (64kbps, HSCSD) data transmission. Intensive expansion of databank and data managing services requires the further increaser of this speed consequently now; due to this progress it is already possible to send IP based packages (up to theoretical limit of 164kbps, GPRS). By this improvement the notion of packet switched data transmission has become known in connection with mobile communication systems. The roaming function has been solved with this. The clients can now use the services of her providers and services have become available even in abroad in about the same quality and range, too. As a result the PAN-European mobile network was created.

The most important element of this period is the idea that mobile communication should reach everyone.

Third generation mobile systems (Abu-El-Ata, 2000) (e.g. UMTS) include the communication systems below.

- Cordless radio systems
- Cellular radio systems
- Satellite radio systems
- Paging
- Private mobile radio system

According to plans the UMTS (Universal Mobile Telecommunications System) was constructed to fulfill such requirements while the speed of the terminals can reach the value of 500 km/h. The frequency bands reserved for the first and second generation is found about at the border of 1 GHz. In order to increase the number of users and to enhance the average data speed (384 kbps) required by the users the set of the UMTS band will be about 2 GHz and the bandwidth will also be extended. The higher data transmission speed opens new prospects in mobile telecommunication. The technical obstacles have considerably been averted from the expansion and development of mobile implementation and applications up to now the developers created one kind of application on a too strong platform and every body was forced to use this application, because there was no other possibility.

However from now onward that system engineers worldwide can supply the customers with always new applications and the customers can decide if they need the offered services or they could like to choose something else. This process could be compared to the PC period of the nineties; when computers of medium size capacity (i386) were developed. Some firm's developed huge quantity of software and the users only had to choose the most suitable of them.

On the field of mobile communication, the problem of the constant but not unlimited bandwidth will give stimulation to the application developers to produce more efficient applications at least till 2012, when fourth generation mobile systems will be able to ensure extremely big bandwidth. According to the plan of Ericsson (Casal, et al, 1999) by 2011 the mobile connection will be equal to an internet access of 100 Mbps.

The third generation which has become a heroic-age has been the period of the group applications approved by customers. It will be the time that the demand for users to transmit and receive their data messages and information in a multimedia environment.

The most intensive stimulus and determining base will be the Internet and its effect will appear in mobile communication systems. Specific combination of test, graphs, animations, voice and real time video broadcast will be usual part of everyday life. The customers will have the intention to reach these facilities also with mobile access equipment.

The packet switched data transmission method, the 3G is IP based (Schweighofer, 1999), consequently there will be no difficulties to embed to the internet-based application in the mobile environment. The 3G systems – even through their construction – follow a philosophy of standardization (Richardson, 2000). UMTS is one of the members of IMT 2000 family which was installed in USA and Japan. These systems are not the same but the modulation type, the allocated bandwidth and the basic in these systems are similar to each other. However, the UMTS has not been able to cover all kind of situation and hence her wireless access networks have right for existing (WLAN, Bluetooth, DECT). By the time of 4G, access to internet through wire or wireless will be no distinguishing factor; the user can recognize a uniform, world-wide communication system in everyday work (Zeng, et al, 2000).

USERS' EXPECTATIONS FROM 4G MOBILE NETWORKS

Presumably, the signs of information society will appear by 2008. Many people will be schooled on high scale so they will be able to apply and to value new services, devices and possibilities. Because of existence of demand, the information market will rise to a significant size and the communication will become a kind of citizen right. Globalization will show up strongly because the everyday life will be simpler and cheaper. An ordinary man see around his world will embrace some scenarios. These potential scenarios (Hjelm, 2000) were contrasted and common points are discussed in the foregoing.

- New type input/output devices will come up for the sake of fast data exchange (glasses displaying 3D virtual world, collapsible screens, e-paper, voice and handwriting recognition).
- New type semiconductor industry will rise (by means of plastic based chip technology the extremely cheap or throw-away electronic tools will be common; 4G terminals will be available for everyone).
- Access to the fourth generation mobile systems will be low-priced (advertisements what could be displayed on the screen of 4G terminals means incomings and users should partake for it).
- Amount of users will reach a high level.
- There will be a heavy competition between applications and service-providers for users.
- Quality of Internet access by wire or wireless will be equal or almost the same (quality of content-providing will be excellent using a mobile terminal).
- Multimedia will be required to the trivial work (multimedia mean a kind of extra information).
- Some economic, social or state groups could maintain own part-networks (virtual private networks will be used well at administration, personal data-managing – for example mobile ID – and voting a president).
- It follows that the mobile networks should be stable and dependable, should be available for 24 hour per day.
- Conception of a global telecommunication system becomes real; for example a telephone or data call from a jungle to an advanced mega city should be trouble free (there are ground settled wired or radio-based backbone network in well built-up areas and anywhere else are satellite-based backbone telecommunication systems).
- Easy interconnection of different systems (e.g. GPS, Internet, other communication networks).

- This effects that the man is not able to vanish in the Earth, but this man could be found easily anywhere he stays.

TECHNICAL CONDITIONS FOR REALIZATION OF FOURTH GENERATION MOBILE NETWORKS

Anatomizing the 4G mobile systems by developing parameters will be a complete network if a set of features are realized like below:-

- Majority of people can access to voice-or data-based services what are provided by mobile networks (This requires efficient resource-management, for example usage of ad hoc extension in wireless systems).
- The mobile network is able to attach to Internet fully because of basic concept of it (in this way LP based technologies would be used through mobile network (e.g. VolP, voice over IP).
- Problem of virtual private networks is worked out their security and authentication technology are improved well).
- The network is able to realign itself (i.e. to manage several type backbone and it use the best one, it means adaptation).

- The system is able to keep on QoS parameters (Quality of Service).
- Parameter of availability of communication network must be close to 100%
- Application of what are required by a daily normal lifestyle will be run on mobile terminals without any restriction (e.g. news reading with multimedia, sending some orders, voice recording, pocket-secretary functions).
- A universal software/hardware interface could be standardized what should facilitate to develop new services without any problem (Easy to develop for four generation mobile systems).

CURRENT TRENDS FOR THE FOURTH GENERATION MOBILE NETWORKS

There are four technical trends (Butach, 2000), reckoned among pioneers in this moment but they have well-grounded concepts. They are: content provision and agents, software radio, managing ad hoc networks and virtual private networks.

In the future, the group of network supporters and the group of content-providers will be different significantly. As at present day we difference between the Internets – i.e. supporters and the information and entertainment providers (e.g. CNN on the Internet). This will permeate into the market of mobile communication.

The customer could pick and choose among the mobile application and their providers which one offers a high level of performance. The user chooses the best newsreader program for their terminal and subsequently the network supporter. The so-called agent – technology (Zeng, et al, 2000) will help in decision making.

Agent means a special, intelligent program which runs basically on the owner terminal or starts his work on it. Its aim is to collect the user’s habits in telecommunication and everyday lifetime and to offer gathered services based on them. Often these programs create new agents of what runs independently from user’s terminal and travel in different networks to complete the tasks of their parent agents. They contact the network supporters and information or application providers to collect information with interest of user in view (Casal, et al, 1999). Finally they compose a list of found and suitable services.

Advantages of this technology are: the agents could accomplish multiple tasks (data mining, self location, finding other agents and users, etc) due to programmed intelligent, the agents usually do not run on user's equipment so they do not allocate its resources but they run on specified systems (for example internet) the processed results could be displayed on the screen of mobile device easily (Fuhua and Korba, 2000).

Software Radio

Software radio (Mitola, 1995) is a newer technology to reduce the number of components of hardware but to keep on the performance of functionality of the radio terminal. Software radio is an emerging technology, thought to build flexible radio systems, multi-user-vice, multi-standard, reconfigurable and reprogrammable by software. The flexibility of a software radio system consists in its capability to operate in multi-service environments, without being constrained to a particular standard, but able to offer, in theory, services of any already standardized systems or future ones on any radio frequency band.

The compatibility of a software radio system with any defined radio mobile is guaranteed by its re-configurability, i.e. by digital signal processing (DSP) engine reprogram ability, which, in real time will implement radio interface and upper layer protocols. The users could take a trip round the world with a software radio, and he is able to communicate anywhere even though the local radio parameters (modulation, bandwidth) could be different from those at home system.

Any changes in functionality of radio could be caused by reloading the software of radio across the air interface (Mitola, 1995).

The software radio concept has several advantages. For manufacturers, there is the possibility to concentrate research and development efforts on a reduced hardware platform set, applicable to every cellular system and market. Mass production of this kind of terminals would allow lowered costs. Operators will be able to rapidly roll out new services tailored towards the needs of each user. The advantages for users are the possibility to roam their communications to other cellular systems and take advantage of the worldwide mobility. Moreover, users can configure their terminals according to their preferences.

Ad Hoc Networks

There is an assumption that the fourth generation mobile networks will mean global sized and uniform systems. In the case of existence of more than one million users the efficient resource managing these networks will be required because the frequency bandwidth and the data processing capabilities are limited. Situations could arise where users are close to the each other but a distant base control processes their radio signals in conformity with current protocols.

A reasonable solution is that the terminals should be able to transact their packet-based traffic among them while there is no need for administration of distance base control. In this way a lot of resources are freed and they could be reallocated. The definition of ad hoc networks (Royer, et al, 2000) as a set of uniform devices where the terminals have the

self same hardware platform and there are no other assigned fix-settled equipments to control the systems is very important. This type of network could provide better performance in perculiar cases.
Situations of ad hoc networks will be usually in 2008 in everyday life. By that time the so-called intelligent household utensils will spread and they will be very popular. In our home the conventional tools and utensils will be “clever”. For example the refrigerator will be able to sense the decrease of foods and it will order them using the Internet; our coffee maker will make an offer about our breakfast; our hearing system will adapt to our present at home; TV will find our programme-watching habits out and it will try to propose to us the available best TV channels.

There are several situations where the information-sources could be connected by ad hoc radio networks. If a new device is brought home it must be put down on the floor, while the other utensils will be put into operation. There is no need for a special center and there is no point to in setting one by automating our home-system. However, there would be outdoor cases (disasters, extended network failures, etc) where the fix settled infrastructures come to unusable states but the communications are required strongly. In this situation a fourth generation network with ad hoc extension could stand in the gap.

VIRTUAL PRIVATE NETWORKS

There are possibilities that the terminal of a world-wide communication the rule that terminals are far cheaper than others. Let’s follows users’ requirement that they want to establish a special private network based on this. For example, a firm would like to deploy a mobile communication system to solve the communication problems between the employees but there are difficulties because of the high price in a complex administrative setup.
Instead, the solution would be easier if the firm could make a contract with a local or a global network supporter to divide its network reallocating resources; so that the network supporter assigns a virtual private network to this firm and workers could access to the network of their firm using a standard device. Of course, the fourth generation mobile networks require high security level, so efficient security and authentication processes should have been worked out. Naturally, stability of these systems are very important, too.

Also bigger social group could have resort to a kind of networks we have to analyze the fourth generation networks whether they will have strong and efficient security processes to guarantee a level of data safety where general civic administration tasks would be done (e.g. prolongation of a driving license or execution of a general election). So practice of civic right would be very cheap and the social decisions would be faster than nowadays.

Disadvantage of this is the dependence upon the technology. Failure of communication network could indicate unforeseeable economic and confidential crisis. As a matter of fact, there will be some reserve forces that will want to play a lone hand (for example military). It is easy to see that kind of extension of fourth generation mobile networks will be developed by above a limit of user requirements.

FACILITIES FOR FIFTH GENERATION SYSTEMS

What kind of network will the fifth be? It is a difficult question. In the future the mankind will conquer the Outer space. Many space stations will circle round the Earth as nowadays Alpha (International Space Station) does and there will be a lot of people to research new technologies, to do industrial works or to take a cosmic sightseeing. This people would like to communicate and the communication market will find a good answer for their requirements. The fifth generation mobile system will be very important in the inter-planet communication.

CONCLUSION

The different direction of the development of mobile networks from the first phase down to the fifth was described in this paper. The first generation created the mobile communication but it had several shortcomings like roaming. The second one brought solutions to most of these problems; it also made it possible to build-up national and continental level networks. It is possible for the Internet related mobile communication networks to be accomplished by UMTS system in Nigeria. Their main characteristics are IP based communication and data transmission, all put into the multimedia environment. Although it is characterized by standards, which tend to be unified, and beside them also are other wireless access networks, which have the reason for existence.

Nowadays there are trends and tendencies in the set requirements of technology and that could not be accomplished by the third generation systems. The fourth generation will manage the global sized user population, and this could meet new human-machine interfaces. In everyday lifestyle a mobile terminal will be indispensable. Naturally, fusion of Internet and mobile telecommunication started in the third generation, but this process will accomplish its purpose in the fourth generation mobile networks.

REFERENCES

[1] **Abu El-Ata, M.A**. (2000) "Evolution of mobile cellular communication systems. The Journey to UMTS". Radio Science Conference, 2000 17th NRSC 2000 Seventeenth National.

[2] **Richardson, K.W**. (2000) " UMTS Overview", Electronics & Communication Engineering Journal, Volume: 12 Issue: 3, June.

[3] **Fan, L, Sheriff, R. E, Gardiner, J.G**.(2000) " Satellite-UMTS service provision using Ip-based technology". Vehicular Technology Conference Proceedings, 2000. VTC 2000 – Spring Tokyo. 2000 IEEE 51st, Volume: 3.

[4] **Hjelm, B.** (2000) " Standards and intellectual property rights – the international standardization of third-generation mobile system", Computers and Communications, 2000 – Proceedings. ISCC 2000. Fifth IEEE symposium.

[5] **Zeng, M, Annamalai, A. Bhargava, V.K.** (2000) " A harmonization of global third-generation mobile systems". IEEE Communication Magazine, Volume: 38 Issue: 12, December.

[6] **Casal, C.R. Schoute, F. Prasad, R**. (1999) "A novel concept for fourth generation mobile multimedia communication" Vehicular Technology Conference 1999. VTC 1999 – Fall IEEE VTS 50th, Volume: 1.

[7] **Schweighofer, E. Merkl, D**. (1999) "Data mining in law with adaptive learning techniques", Database and Expert systems Applications, 1999. Proceedings, Tenth International Workshop, 1999.

[8] **Fuhua Lin; Korba, L**. (2000) " Incorporating communication monitoring and control facility in multi-agent systems", systems, Man and Cybernetics, 2000 IEEE International Conference, Volume 4.

[9] **Butach, E.** (2000) "The software Radio Concept", IEEE Communications Magazine, September, 2000

[10] **Mitola, J.** (1995) "The Software Radio architecture", IEEE Communications Magazine, May 1995.

[11] **Royer, E.M, Chai-Keong T.** (1999) "A review of current routing protocols for ad hoc. Mobile wireless networks", IEEE Personal Communications, Volume ; 6 Issue: 2, April

[12] **Shoewu, O. (2003) "**A Critical Survey of the Short Message service within GSM" Paper Presentation at the National Conference of the Electrical Division, of the Nigerian Society of Engineers 2nd October – 3rd October , 2003

ANALYSIS OF CAPACITY LIMITATION IN NIGERIAN GSM NETWORKS AND THE EFFECTS ON SERVICE PROVIDERS AND SUBSCRIBERS

V.O. MATTHEWS AND ADETIBA E.

Department of Electrical and Information Engineering
Covenant University, Ota, Nigeria
vomatthews@yahoo.com and e_adetiba@yahoo.com

ABSTRACT

The performance of GSM network is measured in terms of KPIs (Key Performance Indicators) based on statistics generated from the network. The most important of these performance indicators from the operators' perspective are BER (bit error rate), the FER (frame error rate) and the DCR (dropped call rate).
The Dropped Call Rate (DCR) is a measure of the calls dropped in a network as it gives a quick overview of network quality and revenues lost. This makes it one of the most important parameters in network optimization. At the frame level in the NMS (Network Management System), the DCR is measured against the Slow Associated Control Channel (SACCH) frame. If the SACCH frame is not received, then it is considered to be dropped calls.
For this work data was acquired form the Network Management System of various GSM operators in Nigeria (e.g. MTN, Celtel, Globacom etc.). The acquired data was analyzed to statistically illustrate the extent of revenue that is lost as a result of dropped calls and the consequent impact on the customers/subscribers.

Keywords: GSM, KPIs, BER, FER, DCR, SACCH etc

INTRODUCTION

Quality degradation in a GSM network can be due to many factors such as capacity limitation, interference, unfavorable propagation conditions, blocking and etc.
Without adequate capacity, users will not be able to enter the network even though there might be suitable coverage in the area. Therefore providing the correct capacity in the correct location at the correct time is essential to maximize revenue generation, ensure high utilization of installed network infrastructure and provide seamless service for subscribers' satisfaction. However, capacity, both in terms of bandwidth and hardware will always be limited in a practical communication system and the primary method that is employed in arresting capacity limitation is called Handover. This involves relaying of the excess traffic to a less congested neighboring cell and it is only efficient if there is a major population drift from such cell to the congested cell.

Theoretical Analysis of Capacity Limitation in Nigerian GSM Networks

Due to the limited bandwidth that is available to GSM operators, heavy traffic is usually generated during special events at some cell sites (e.g. religious camps, stadia, markets etc) and the traffic sometimes exceeds the bandwidth capacity of the assigned base station.

Based on this, a high dropped calls which results in huge lost of revenue by the operators and erratic service provision to the customers is regularly experienced.
The bandwidth allotted to GSM operators in Nigeria is as shown below.

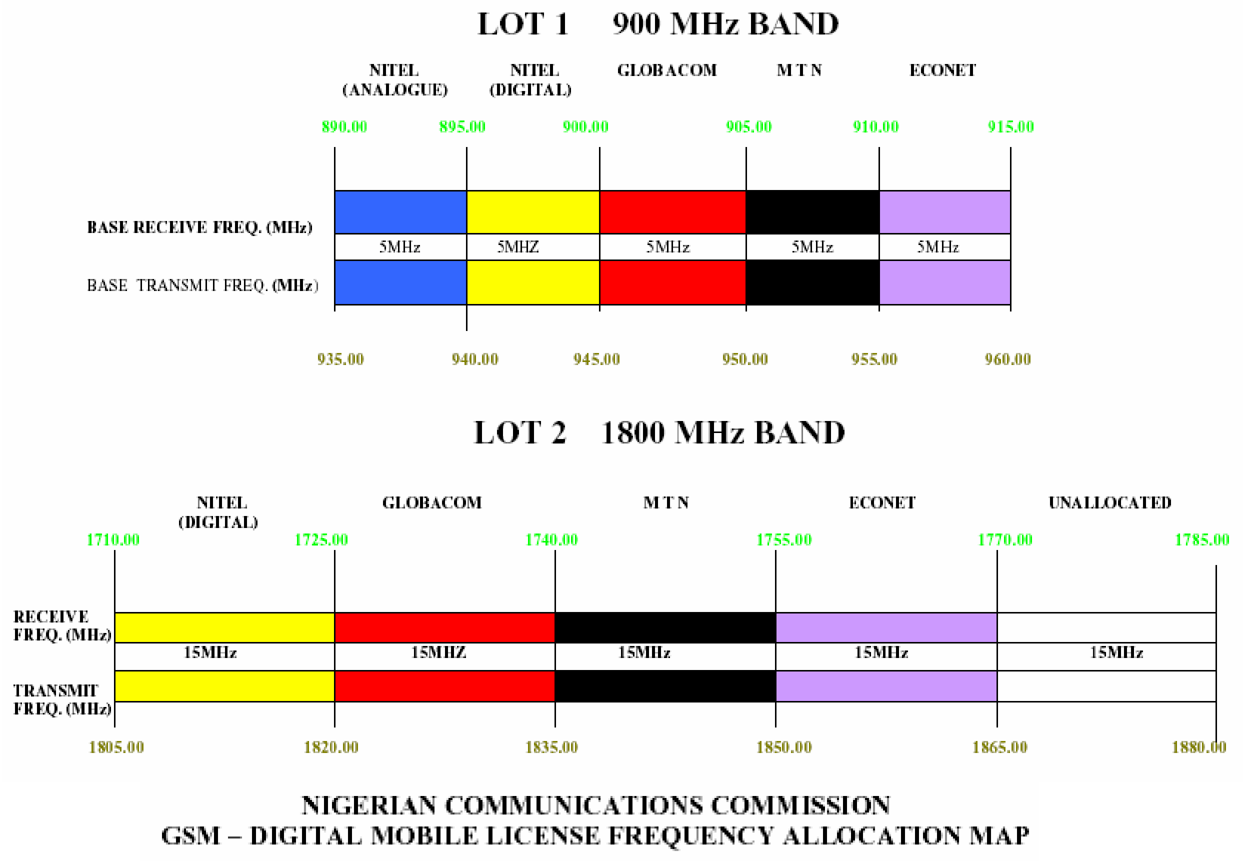


Fig 1.0 (NCC Website)

In MTN network for instance, Canaan Land is served by site T325. This has 2 cells i.e. one 900 band cell (325A) and one 1800 band cell (325 DA).

325A (the 900 band cell on Canaan land) has; - 6 Transceivers (TRX) - 42 Traffic Channels (TCh - for both Voice and Data) - 5 Stand Alone Dedicated Control Channels (SDCCh - for call setup, sms and location update) – 1 Broadcast Control Channel (BCCh - for broadcast by BTS to show network attachment)

While 325DA (the 1800 band cell) has; - 6 Transceivers (TRX) - 45 TCh - 2 SDCCh - 1 BCCh

Similarly, Redemption Camp is served by site T71. This has 6 cells i.e. three 900 band cells (71A, 71B and 71C) and three 1800 band cells (71DA, 71DB and 71DC).

71A has - 4TRX - 29 TCh - 2 SDCCh - 1 BCCh

71B has - 4 TRX - 27 TCh - 4 SDCCh - 1 BCCh

71C has - 4 TRX - 27 TCh - 4 SDCCh - 1 BCCh

71DA has - 6 TRX - 45 TCh - 2 SDCCh - 1 BCCh

71DB has - 6 TRX - 45 TCh - 2 SDCCh - 1 BCCh

71DC has - 6 TRX - 45 TCh - 2 SDCCh - 1 BCCh

The maximum number of users varies according to the size of the cell i.e. number of configured Traffic Channels. MTN for instance uses a 2% grade of service (GoS) and Erlang is the unit of traffic used in telecommunications, which is measured as the total traffic volume of a cell in 1hr.

i.e. Erlang = no of calls x duration/60mins (1hr) ---------------(1)

From the above information, it is expected that 325A (the 900 band cell on Canaan land) with 42 TCh should have an Erlang of 32.84 at peak period, using a GoS of 2% from Erlang B Table. Hence, the number of calls or callers/users can be calculated from the formula above which gives us 394 users assuming an average of 5mins call duration.

Also, 325DA (the 1800 band cell on Canaan land) with 45 TCh has an Erlang of 35.61 at peak period. The maximum number of callers in this cell, assuming an average of 5 mins call duration is 427 callers/users.

Therefore, the total maximum number of users on Canaan land during peak periods assuming call duration of 5 mins is:

394 + 427 = 821 callers/users. --------------------- (2) [3, 4, 8, 11]

It should be noted that any caller beyond the maximum value calculated above for each base station is automatically dropped. Using the formula in equation 1.0 above, Celtel cell parameters at some designated cell sites are shown in Table 1.0 below.

SITES	**CELLS**	**SECTOR**	**No of TCh**	**FR Traffic (Erlang)**	**HR Traffic (Erlang)**	**NO OF FR USER**	**NO OF HR USER**
CANNAN LAND	LG414D1	1	46	36.53	54.795	1461	2191.8
	LG414D2	2	46	36.53	54.795	1461	2191.8
	LG414D3	3	46	36.53	54.795	1461	2191.8
	LG414G1	1	14	8.2002	12.3003	328	492.012
	LG414G2	2	14	8.2002	12.3003	328	492.012
	LG414G3	3	14	8.2002	12.3003	328	492.012
REDEMPTION CAMP	LI048D1	1	14	8.2002	12.3003	328	492.012
	LI048D2	2	14	8.2002	12.3003	328	492.012

	LI048D3	3	14	8.2002	12.3003	328	492.012
	LI048G1	1	14	8.2002	12.3003	328	492.012
	LI048G2	2	14	8.2002	12.3003	328	492.012
	LI048G3	3	13	7.402	11.103	296	444.12
ABEOKUTA STADIUM	ABY01D1	1	14	8.2002	12.3003	328	492.012
	ABY01D2	2	14	8.2002	12.3003	328	492.012
	ABY01D3	3	14	8.2002	12.3003	328	492.012
	ABY01G1	1	13	7.402	11.103	296	444.12
	ABY01G2	2	13	7.402	11.103	296	444.12
	ABY01G3	3	13	7.402	11.103	296	444.12
NATIONAL STADIUM	LG725D1	1	30	21.93	32.895	877	1315.8
	LG725D2	2	30	21.93	32.895	877	1315.8
	LG725D3	3	30	21.93	32.895	877	1315.8
	LG128D1	1	45	35.62	53.43	1425	2137.2
	LG128D2	2	46	36.53	54.795	1461	2191.8
	LG128D3	3	30	21.93	32.895	877	1315.8
	LG128G1	1	14	8.2002	12.3003	328	492.012
	LG128G2	2	13	7.402	11.103	296	444.12
	LG128G3	3	14	8.2002	12.3003	328	492.012

Table 1.0: Celtel Cell Parameters for Some Sites

(Source: Celtel Sites Documentation)

Statistics of Dropped Calls/Network Congestion on Canaan land Cell Site

The following empirical data (Table 2.0) from Celtel Network Management System (NMS) for weekly call statistics at **Celtel** cells on Canaan Land further illustrates the traffic congestion problem when a cell is overpopulated.

Sector OBJECTID	DATE	HO	DROPS	DROP_CALL_RATE
Sector 5	11/14/2005	0	16	2.48

Sector 1	11/14/2005	0	57	9.74
Sector 3	11/14/2005	0	75	10.37
Sector 2	11/14/2005	0	24	5.37
Sector 6	11/14/2005	0	20	3.21
Sector 4	11/14/2005	0	22	2.84
Sector 2	11/21/2005	0	129	2.45
Sector 1	11/21/2005	0	269	2.96
Sector 5	11/21/2005	0	32	1.25
Sector 3	11/21/2005	0	233	3.7
Sector 6	11/21/2005	0	58	1.14
Sector 4	11/21/2005	0	52	1.28
Sector 5	11/28/2005	0	163	1.02
Sector 1	11/28/2005	0	1428	2.35
Sector 4	11/28/2005	0	207	0.87
Sector 3	11/28/2005	0	2120	2.9
Sector 6	11/28/2005	0	263	1.11
Sector 2	11/28/2005	0	1304	2.82
Sector 3	**12/5/2005**	**0**	**1993**	**1.56**
Sector 4	**12/5/2005**	**0**	**2099**	**0.78**
Sector 6	**12/5/2005**	**0**	**303**	**0.8**
Sector 1	**12/5/2005**	**0**	**3825**	**1.23**
Sector 5	**12/5/2005**	**0**	**1110**	**0.6**
Sector 2	**12/5/2005**	**0**	**2492**	**1.15**
Sector 1	12/12/2005	0	1184	2.1
Sector 3	12/12/2005	0	1914	2.41
Sector 2	12/12/2005	0	785	2.32
Sector 6	12/12/2005	0	132	1.14
Sector 5	12/12/2005	0	72	1.25
Sector 4	12/12/2005	0	113	1.02
Sector 6	12/19/2005	0	168	1.21
Sector 5	12/19/2005	0	98	1.81
Sector 2	12/19/2005	0	1095	2.43

Sector 3	12/19/2005	0	2613	2.53
Sector 4	12/19/2005	0	92	0.68
Sector 1	12/19/2005	0	1601	2.24
Sector 4	12/26/2005	0	124	1.13

Table 2.0: Celtel NMS call statistics from 14/11/2005 to 30/01/2006

(Source: Celtel Network Management System)

It can be seen on the table that on the first day of Shiloh 2005 i.e. (5/12/2005), there was a consistent high dropped calls in all the sectors. This is due to the sudden surge of population on Canaan land for the annual Shiloh event.

For further illustration, figure 4.0 below shows the tabulated weekly dropped calls at Canaan land in December 2005 by MTN network on a bar chart. Peak dropped calls were experienced during Shiloh week as earlier predicted.

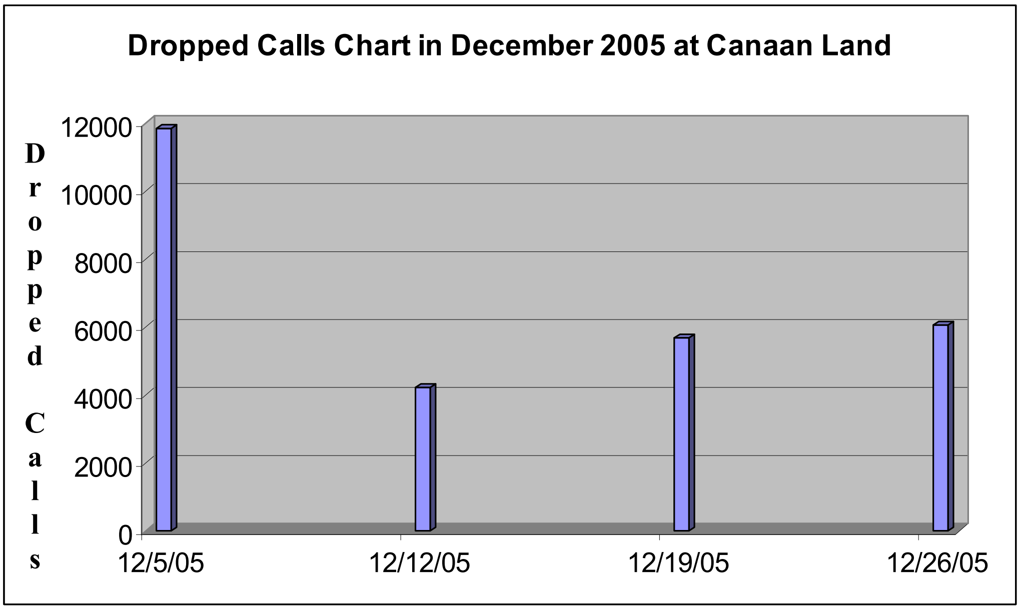


Fig 2.0 Dropped calls at Canaan Land in December 2005 (MTN Network)

Logically, the lost of revenue incurred by cellular operators is in direct proportion to the dropped calls over a given period.

Mitigation of GSM Network Capacity Limitation

To increase capacity given a limited bandwidth, **Frequency Re - use** is implemented by radio frequency engineers. In Nigeria for instance, Frequency Re - use is a way of planning the re -use of the frequencies assigned to every operator by NCC (Nigerian Communications Commission). For example the maximum available frequencies on 1800 band in Nigeria must be shared amongst MTN, Celtel, Globacom and M-Tel, hence the need for re -use. The pattern employed by MTN is 4 X 3 i.e. re -use frequencies after every 4 sites with 3 cells each. Meanwhile, in this scheme, there is no consideration to cater for traffic congestion due to overpopulation.

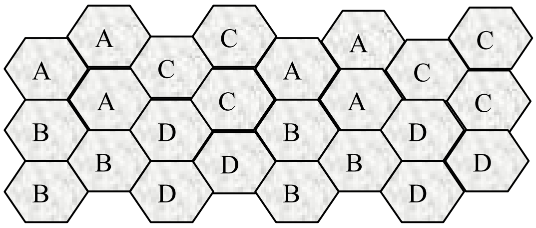


Fig 2.0 MTN (Nigeria) 4 X 3 Frequency Reuse Pattern.

To increase capacity in some cellular networks where the base station capacity is limited by the number of transceivers, **a 'bunch' concept consisting of a central unit (CU) and remote antenna units (RAUs) was proposed for UMTS**.

A typical bunched cellular network is shown in fig 3.0 below.

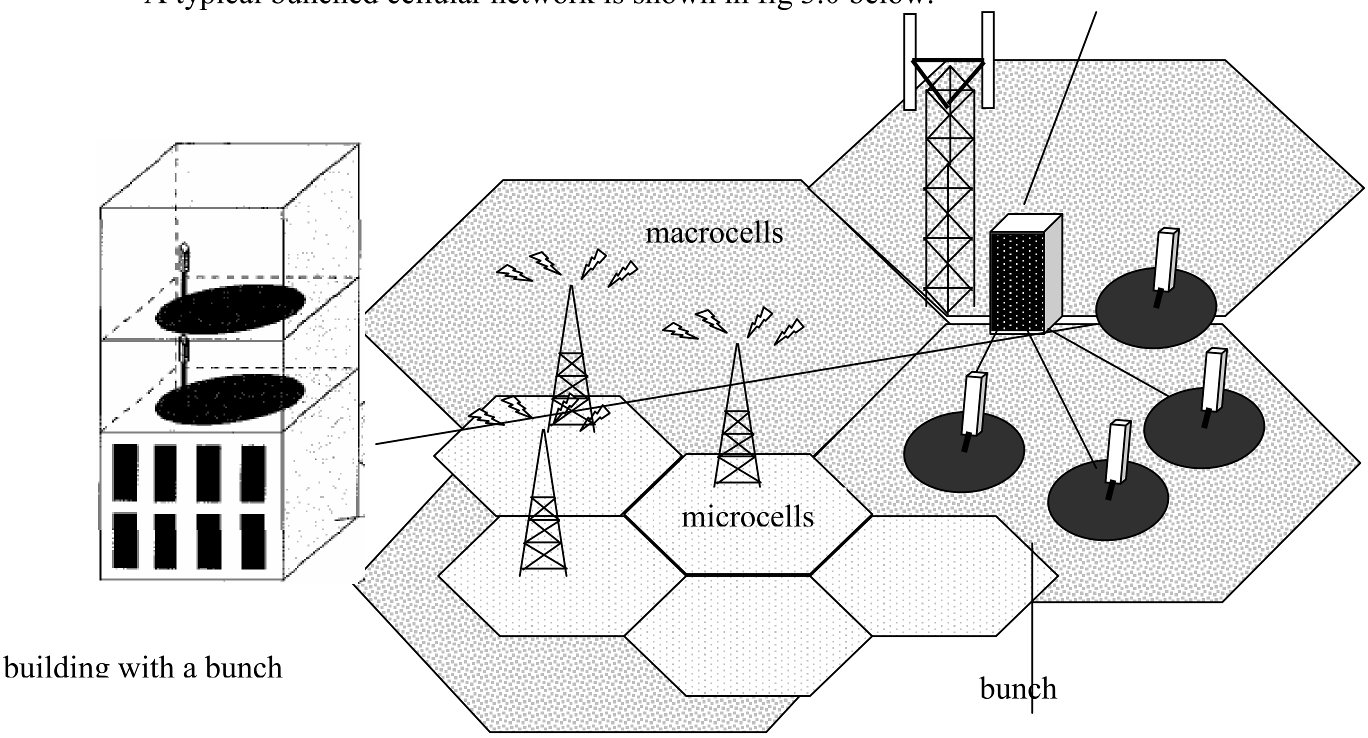


Fig 3.0 A Scenario of mixed cell structure with Bunched Network

The bunch concept involves a Central Unit (CU) that controls a set of remote antennas or base stations. The Central Units deal with all decisions on channel allocation, service request and handover. The concept can be viewed simply as a very advanced base station with a number of small antennas for remote sensing. The central unit will therefore have complete control over all the traffic in its coverage area and will be able to maximize the resource utilization for the current traffic. This provides opportunities for uplink diversity and avoids intercell handovers in its coverage area. [1, 11, 13, 15, 17, 22].

Intelligent Relaying (IR) is another technique used by cellular telecommunication operators to minimize the amount of planning and the number of base stations required in a cellular network and to improve cell capacity. A network employing intelligent relaying includes mobiles that are capable of passing data directly to other mobiles, rather than directly to a base station, as for a conventional cellular network [1].

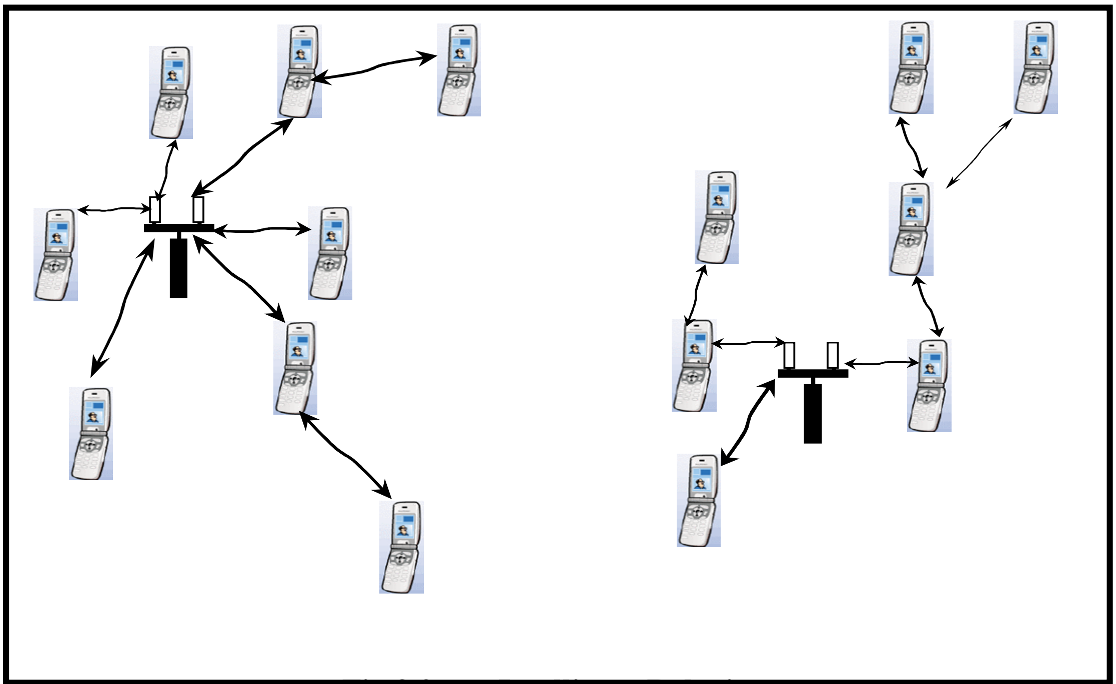


Fig 3.0 Intelligent Relaying

In IR, mobile terminals are permitted to receive and retransmit data on behalf of other users. A cellular network that implements IR is shown in fig 3.0 above which shows how the route between mobile and base station can be broken down into shorter mobile-mobile hops and how IR nodes can act as forwarding nodes for a number of mobiles.
To plan a network incorporating IR for capacity enhancement, each mobile is considered as a "virtual cell", acting as a base station at its center. The coverage area of the virtual cell varies with respect to factors like; change in the transmit power of the mobile, mobility of the user and the number of other mobiles in the vicinity that are available to relay data. Where IR is implemented as a network option, the virtual cells can be considered as an overlay to a conventional cell structure.[1, 7, 8, 16]
Several other approaches such as **dynamic cell sizing** and **radio resource management** are employed to mitigate capacity limitation in a GSM Network and any or combinations of these approaches could be implemented as appropriate.

Conclusion and Recommendation

This paper has analyzed the problem of network congestion due to capacity limitations which usually results in reduction of quality of service (QOS) in cellular communications system. Since QOS is an essential performance criterion, cost effective solutions to minimize the problem and boost the network performance have been explored.
However to further mitigate this capacity limitation problem, we are currently working on a Software radio based approach, to design and simulate a capacity enhancement Nomadic Base Transceiver Station Sub-System (NBTSS).

References

[1] A.G.SPILLING et al (2000), “Self-organization in future mobile communications”, *Electronic and Communication Engineering Journal*, June 2000, pp.133-146.

[2] BAMBOS, N. (1998), “Toward power-sensitive network architectures in wireless communications: concepts, issues and design”, *IEEE Personal.Commununication,* June 1998, pp.50-59.

[3] KATZELA, I. and NAGHSHINEM, M (1996), “Channel Assignment schemes for cellular mobile telecommunication systems: a comprehensive survey”, *IEEE Personal Communication,* June 1996, pp 10-31.

[4] THEODORE S.RAPPAPORT (1996), Wireless Communications Principles and Practice, Publishing House of Electronics Industry, Beijing.

[5] BO BERGLUND et al (2005), “Third design release of Ericsson’s WCDMA macro radio base stations”, Ericsson Review No 2, 2005, pp. 70-82.

[6] ERICSSON “Base Station System (BSS) Architecture”, GSM RBS 2000 Baiscs, pp. 6-71.

[7] ERICSSON “Base Station Controller and Transcoder Controller”, GSM System Survey, pp. 133-153.

[8] AIRCOM (2002), “GSM System Overview”, GSM System Overview Aircom International Grovesnor House, 2002, pp. 1-162.

[9] JEREMY BRADBURY (2000), Linear Predictive Coding, www.lpcsynthesis.com, pp. 1-23.

[10] JUSTIN CHEN (2004), “Voice Conversion Experiment and Conclusion”, Connexions module:m12478, Version 1.4, US/Central 2004, pp. 1-2

[11] MACARIO R.C.V. (1997), Cellular Radio Principles and Design, Macmillan Press Ltd, London.

[12] MICHEL DAOUD YACOUB (2002), WIRELESS TECHNOLOGY Protocol, Standards and Techniques, CRC PRESS, United States of America.

[13] ERICSSON (1998), GSM RBS 2000 BASICS, STUDENT TEXT EN/LZT 123 3805 R2A.

[14] DAG STRANNEBY WILLIAM WALKER (2004), Digital Signal Processing and Application, Newnes Linacre House, Jordan Hill, Oxford OX2 8DP.

[15] BOUCHER, N.J.(1990). Cellular Radio Handbook, Quantum Publishing Inc, USA.

[16] LEE,W.C.Y.(1989). Mobile Cellular Communications Systems, McGraw-Hill Book Co; NY

[17] YOUNG, W.R. (1979), “Advance mobile phone services: Introduction, background and objective”, B.S.T.J., 58, pp1-41.

[18] BALSTON, D.M. and MACARIO R.C.V. (1993). Cellular Radio Systems, Artech House.

[19] GARDINER, J. and WEST, B(1995). Personal Communications Systems and Technologies, Artech House.

[20] GIBSON, D.(Ed) (1996). The Mobile Communication Handbook, CRC Press, in co-operation with IEEE.
[21] LEE, C.Y. (1992). Mobile Communications Design Fundamentals, Second Edition, Wiley.
[22] MACARIO, R.C.V.(Ed) (1996). Mobile Personal Radio Systems, The IEE Press.
[23] MEHROTRA, A. (1994). Cellular Radio Analog and Digital Systems, Artech House.
[24] MEHROTRA, A. (1994). Cellular Radio Performance Engineering, Artech House.
[25] MOULY, M. and PAUTET, M. (1992). The GSM System for Mobile Communications. Published by the authors, Paris.
[26] STEELE, R.(1992). Mobile Radio Communications, Pentech Press.
[27] WONG, P. and BRITLAND, D. (1995) Mobile Data Communication Systems, Artech House.
[28] GRAFFITS, J.(1987). Radio Wave Propagation and Antennas, Prentice Hall International Ltd, UK.
[29] PARSONS, J.D.(1992). The Mobile Radio Propagation Channel, Pentech Press, London, U.K.
[30] IEEE Vech Tech report on propagation (1988). “Coverage Prediction for mobile radio systems operating in the 800/900 MHz frequency range”, IEEE Veh Tech, VT-37, Feb, pp 3-70.

THE FUTURE OF MOBILE INDUSTRY

R.S. ADESUYI

Department of Electrical & Information Engineering
Covenant University, Ota, Nigeria
e-mail:sundayadesuyi@yahoo.com

Abstract

This paper discusses the future of mobile industry along with some of the background leading to the emergence of wireless technology. First, it gives an overview of today's telecommunication network and the major differences between fixed wired networks and wireless networks. The discussion then focuses on the challenges facing the wireless industry and the way out through aggressive innovation by employing Wireless Intelligence Network (WIN) technology. The paper also discusses some important trends in wireless industry and the customers expectations which are also part of the challenges for the mobile industry. Finally particular reference is made to the developing nations especially Nigeria in the ongoing trends in mobile communication industry.

Introduction
Telecommunications plays a vital role in international commerce and in industrialized nations, it is an accepted necessity. The telecommunication networks in all countries are linked together to form a global telecommunication network for carrying information of all kinds. The Public Switched Telephone Network (PSTN) was originally developed mainly for carrying voice communication but today carries ever increasing data communication traffic. The Internet uses PSTN circuit to carry its data and the rapid growth ot Internet technology has enhanced the growth of data usage in the PSTN.

Wireless or mobile communication services are commonly influencing the growth of telecommunication. On the other hand, in developing countries, wireless communication enables many customers in the cities to obtain affordable telecommunication services. However, it is a matter of fact that during the last fifteen years mobile communications have been a major contributor to the growth of the telecommunications industry as a whole and one of the most dynamic factor of world's economy [8].

Wireless was the original word for "radio" essentially it meant telegraphy. The term wireless was widely used in early twentieth century but dropped out of usage in many parts of the world and was almost replaced by the word radio before the World War II.[2]

Today however, the term wireless is back in a more sophisticated way. Although, it still means "radio", it now has some specific modern applications.
The contemporary use of the word wireless is in describing enormous cellular telephone industry or mobile industry. After all, cell phones are basically sophisticated two-way radios.

Other technologies that involve wireless systems include wireless Local-Area Networks (LANs); Wide Area Networks (WANs); Personal-Area Networks (PANs); Radio Frequency Indentification (RFID); Infrared wireless devices. The latest wireless technology is Ultra Wide Band (UWB), a technology used in wireless networking, mobile computing and low-cost short-range radar system.[2]

WIRELESS COMMUNICATIONS TODAY

The impact of telecommunication on people's lives has been far reaching and is likely to accelerate as wireless technologies pervade the international societies of both developed and developing nations of the world.

Cellular communication technology is the fastest growing and most demanding telecommunications applications. Today, it represent a continuously increasingly percentage of all the new telephone subscriptions around the world.[8]

As a more recent trend makes clear, wireless communications and Internet technologies are converging giving people ever-increasing freedom to choose their methods of communication. This commonality between the Internet and wireless communication is according to many analysts, driving the growth of the wireless industry both regionally, locally, and globally.

As Dr. Irwin Jacobs Chairman and CEO of Qualcomm put it. "The Internet continue to be the fastest growing technology today. What is perhaps less well-known is that wireless is the second fastest growing technology. The true potential of the wireless industry is now being realized as the cost of service drops and wireless networks become more pervasive. Continuing to drive this wireless trend; consumers will access the Internet through a wireless connection. The ability of the next generation systems to effectively transmit voice and data will bring to fruition the convergence of computing and telephoning and with it a whole new range of capabilities and devices[10].

THE WIRELESS REVOLUTION

How fast is the communications revolution moving? Within the next five years stationery desktop systems will no longer be the tool of choice for accessing the Internet. Mobile devices like smart phones and other type of hand-held devices will enable the net to float free of its traditional fixed-wired and provide users wherever they may be with access to e-mail, sports scores; stock quotes, e-banking, flight status, shopping tips; traffic alerts; driving directions and much more[5]

What is making this phenomenon possible is none other than advanced wireless technologies – technologies that are fuelling an incredible expression of voice, as well as data services in every corner of the world. Motorola estimated that by 2005, the number of wireless devices with Internet access will eventually exceed the number of fixed-wired ones[5] . All these predictions has virtually come to pass or they are on the verge of coming to pass.

WIRELESS INDUSTRY DEFINED

According to National Telecommunication and Information Administration (NTIA), telecommunication refers to "any transmission, emission or reception of signs, signals, writing, images and sounds or intelligence (information) of any nature by wire, radio, optical (visual) or other electromagnetic systems." Telecommunications includes satellite communication (wireless but not cellular), cellular wireless (including infrastructure, phones, and PDA – personal digital assistants – chips) land lines (cable), communications equipment (radio and TV), and Internet networks[10].

The wireless industry focused on these devices and services such as satellites services / carriers / operators; mobile communications, pagers and cellular; and personal communications systems and hand held objects / components.

WIRELESS TECHNOLOGY

Parts of the technology of wireless are those that translate voice into digital transmissions. The primary choices worldwide are Qualcomm's CDMA (Code Division Multiple Access). This technology is well adopted in USA and other parts of the world including Europe and Asia. Then we have TDMA (Time Division Multiple Access; Spread Spectrum (SS) which used TDM as the driving technology. We also have GSM (Global Systems for Mobile Communication), which is the dominant player in Europe and had gained wide popularity world wide in terms of subscribers. Other technologies include Bluetooth – which permits communications at short distance; GPRS (General Pocket Radio Services, GPS (Global Position Systems); WAP (Wireless Application Protocol) and other systems that integrate voice / data / video, Internet, and wireless technologies for numerous consumers products.

THE INFLUENCE OF SAN DIEGO ON THE RAPID GROWTH OF MOBILE INDUSTRY

There is no way we can discuss or write about mobile industry without making reference to San Diego, California, US. San Diego has become a reference point in mobile industry worldwide. The region is widely recognized as the "wireless capital of the world." The implications of San Diego, California, United States cannot be overemphasized. As stated by another industry analyst; "In the wireless war for supremacy, San Diego is emerging as the industry boot camp of innovation."[10]. San Diego innovative market made it home to many of the major player in the wireless industry – Qualcomm – (the company that developed the technology of CDMA); Huges; Nokia, Erricson, Motorola, Kyocera and a host of others.

In San Diego, the wireless industry represents a technological sector within the communication chapter similar to the Silicon Valley. In recent years, San Diego has emerged as an international leader in meeting consumer demand – both for wireless technology and equipment. The region's industry has evolved from its early history from University of California, San Diego (UCSD) research to comprehensive sector with significant support organizations.

Insiders trace the beginning of this industry to Linkabit, a 1968 business venture of Professors at the University of California, San Diego, starting in 1980s, Linlabit spun off Qualcomm, and more than 30 others telecommunication companies including Viasat (Satellite communications firms), Wave Ware (specializing in software and connectivity), and Leap Wireless (an operator of wireless networks). These in turn fuelled generators of newer companies, creating a vortex of activity that attracted major international companies now located there.[10]

EVOLUTION OF MOBILE TELEPHONE GSM SYSTMES

Cellular communication technology is the fastest growing and most demanding telecommunications applications. Today it represents a continuously increasing percentage of all be new telephone subscriptions around the world. Currently there are more than 300 million cellular subscribers worldwide[8], and recently 80% of those subscribers are located in United States , Europe,and some parts of Asia.

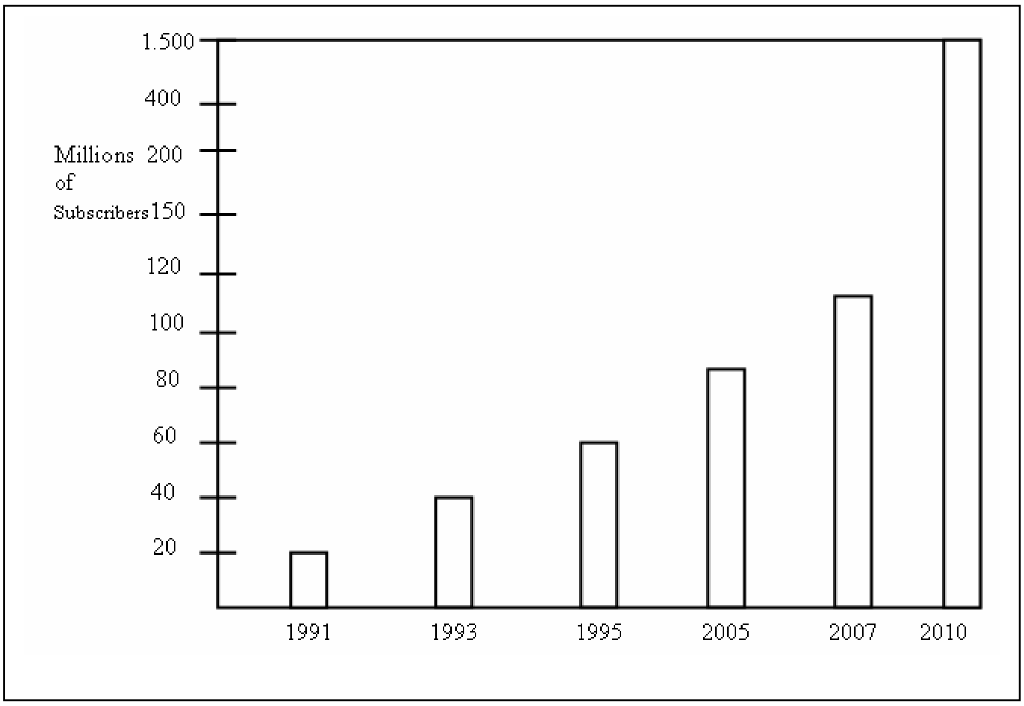
It was predicted that cellular system using digital technology will become the universal method of telecommunication. It has been estimated that some countries may have more mobile phones than fixed wired phones by the year 2000[8]. This estimate has already come to pass in Nigeria because mobile phone has almost entirely replaced fixed-wired telephones[14].

CELLULAR SUBSCRIBER GROWTH WORLDWIDE

The concept of cellular communication service is the use of low – power transmitters where frequencies can be reused within a geographic area. The idea of a cell – based mobile radio service was formulated in the US at the Bell laboratory in the early 1970s.

The Nordic countries were the first to introduce cellular services for commercial use with the introduction of the Nordic mobile telephone (NMT) in 1981. However cellular services or cellular systems began in the United States with the release of the Advanced Mobile Phone Service (AMPS) system in 1983. Later, the AMPS standard was adopted by Asia, later America and the countries located along the Pacific Oceans, creating the largest potential market in the world of cellular[8]

.



Cellular Subscribers Growth Worldwide

ANALOG MOBILE PHONES

In the early 1980, most mobile phones systems were analog rather than digital, like today's newer system. One challenge facing analog system was inability to handle the growing capacity needs in a cost-efficient manner. As a result, digital technology was welcomed.

The advantages of digital system over analog systems include ease of signaling, lower levels of interference, integration of transmission and switching, and increased ability to meet capacity demands.

OVERVIEW AND DEVELOPMENT OF WORLDWIDE MOBILE TELEPHONE SYSTEMS

YEAR	MOBILE SYSTMS
1981	Nordic Mobile Telephone NMT 450
1983	American Mobile Phone System AMPS
1985	Total Access Communication System (TACS) British
1986	Nordic Mobile Telephone (NMT) 900
1991	GSM
1992	Digital cellular system (DCS) 1800
1994	Personal Digital Cellular (PDC)
1998	PCS 1900 – Canada
1996	PCS United States

Table 1: Overview of mobile telephone worldwide   Source: IEC2000

GSM STANDARDISATION

As the cellular communication technology was been evolved, various systems have been developed without the standardized specification. This presented many problems directly related to compatibility and inter-operability, especially with the development of digital radio technology. The GSM standard was intended to address these problems.

From 1982 to 1985, discussions were held to decide between building an analog or digital systems. After a series of field test, a digital system was adopted for GSM. The next task was to decide between narrow or broad band solutions. Therefore in May 1987, the Narrow Band Time division multiple access (TDMA) was adopted[8]. A summary of GSM milestone is given on table 2. bolow.

YEAR	MILESTONE
1982	GSM formed established
1986	Field test were carried out
1987	TDMA -------- as access method
1988	Memorandum of Understanding signed
1989	Validation of GSM System
1990	Preparation system
1991	Commercial system start – up
1992	Coverage of large cities/airport
1993	Coverage of ---
1995	Coverage of rural areas

Table 2: GSM Mile Stone Source: IEC2000

OVERVIEW OF MOBILE PHONE SYSTEMS

Today, there are large number of mobile phones available in the telecommunications market. Whilst not all are in use today, some of the older systems have been superseded and some of the newer systems have not all been rolled out yet, nevertheless many different names and technologies are talked about. The table below gives a summary of the main systems that have been used, are being used or are due for introduction.

Table3.

Cellphone System	Generation	Channel Spacing	Access Method	Comments
AMPS	1G	30 kHz	FDMA	Advanced Mobile Phone System, this analogue first developed and used in the USA
NAMPS	1G	10 kHz	FDMA	Narrow band version of AMPS chiefly used in the USA and Israel based on a 10 kHz channel spacing.
TAC	1G	25 kHz	FDMA	Analogue system originally in the UK. Based around 900 MHz, this system spread world wide. After the system was first introduced, further channels were allocated to reduce congestion, in a standard known as Extended TACS or ETACS.
NMT	1G	12.5 kHz	FDMA	Nordic Mobile Telephone. This analogue system was the first system to be widely used commercially being launched in 1979. It was used initially on 450 MHz and later at 900 MHz. It was used chiefly in Scandinavia but it was adopted by up to 30 other countries including Oman.
NTT	1G	25 kHz	FDMA	Nippon Telegraph and Telephone. The system used in Japan, using a 900 MHz frequency band, and 55 MHz transmit receive spacing. (A high capacity version is known as HICAP).
C450	1G	20 kHz	FDMA	The system adopted in West Germany (East Germany was separate at this time). It used a band in the region of 450 MHz along with a 10 MHz receive / transmit spacing.
GSM	2G	200 kHz	TDMA	Originally called Groupe Speciale Mobile, the initials later stood for Global System for Mobile Communications. It was developed in Europe and first introduced in 1991. The service is normally based around 900 MHz although some 850 MHz allocations exist in the USA.
DCS 1800	2G	200 kHz	TDMA	1800 MHz derivation of GSM and is also known as GSM 1800.
PCS 1900	2G	200 kHz	TDMA	1900 MHz derivation of GSM and is also known as GSM 1900.

TDMA	2G	30 kHz	TDMA	Although it was originally known as US Digital Cellular (USDC) and was introduced in 1991. It is sometimes called North America Digital Cellular and also known by its standard number IS54 that was later updated to standard IS136. It is a 2G digital system that was designed to operate alongside the AMPS system.
PDC	2G	25 kHz	TDMA	Pacific or Personal Digital Cellular. The system found only in Japan where it has gained very widespread use. It has many similarities with IS54 although it uses a different speech coder and a 25 kHz bandwidth.
GPRS	2.5G	200 kHz	TDMA	General Packet Radio Service. A data service that can be layered onto GSM. It uses packet switching instead of circuit switching to provide the required performance. Data rates of up to 115 kbps attainable.
EDGE	2.5 / 3G	200 kHz	TDMA	Enhanced Data rates for GSM Evolution. The system uses s different form of modulation (8PSK) and packet switching, which is overlayed on top of GSM to provide the enhanced performance. Systems using the EDGE system may also be known as EGPRS systems.
CdmaOne	2G	1.25 MHz	CDMA	This is the brand name for the system with the standard reference IS95. It was the first CDMA system to gain widespread use. The initial specification for the system was IS95A, but its performance was later upgraded under IS95B, which the cdmaOne specification actually uses, Apart from voice it also carries data at rates up to 14.4 kbps for IS95A and under IS95B data rate of up to 115 kbps are supported.
CDMA2000 1X	2.5G	1.25 MHz	CDMA	This system supports both voice and data capabilities within a standard 1.25 MHz CDMA channel CDMA2000 builds on cdmaOne to provide an evolution path to 3G. The system doubles the voice capacity of cdmaOne systems and also supports high-speed data services. Peak data rates of 153 kbps are currently achievable with figure of 307 kbps quoted for the future, and 614 kbps when two channels are used.
CDMA2000 1xEV-DO	3G	1.25 MHz	CDMA	The ED-DO stands for Evolution Data Only. This is an evolution of CDMA2000 that is designed for data only use and its specification is IS856. It provides peak data rate capability of over 2.45 Mbps on the forward or downlink, i.e. from the base station to the user. The aim of the system is to deliver a low cost per megabyte capability along with an always on connection costed on the data downloaded rather than connection time.
CDMA2000 EV-DV	3G	1.25 MHz	CDMA	This stands for Evolution Data and Voice. It is an evolution of CDMA2000 that can simultaneously transmit voice and data. The peak data rate is 3.1 Mbps on the forward link. The reverse link is very similar to CDMA2000 and is limited to 384 kbps.
UMTS	3G	5 MHz	CDMA/ TDMA	Universal Mobile Telecommunication System. Uses Wideband CDMA(W-CDMA) with one 5 MHz wide

				channel for both voice and data, providing data speeds up to 2 Mbps.
TD-SCDMA	3G	1.6 MHz	CDMA	Time Division Synchronous CDMA. A system developed in China to establish their position on the cellular telecommunications arena. It uses the same bands for transmit and receive allowing different time slots for base stations and mobiles to communicate. Unlike other 3G systems it uses only a time division duplex (TDD) system..

Table 3: Extracted from Radio-Electronic.Com.

NEED FOR WIRELESS COMMUNICATION

One of the reasons why wireless communication started to develop is mobility –co-ordination. Looking broadly at the way which mobility has been co-ordinated in human communication and interactions, one can see three general phases.
Phase one covers the period before the advent of telegraph, telecommunication regarding mobility could only be delivered by being visible i.e. the speed of the message was the speed of physical transport.[2] .

With the development of Telecommunication, which is the second phase, the speed of the messages that could potentially cause travel were able to move at many times the speed of physical travel. Thus one could send a message to a remote person without the need for a physical messenger. (de Sola, 1980).

The limitation of this systems is that in order for the message to come through, the person sending a message need access to a sending device at a fixed a location, and also needs to know the physical location of the person who is receiving the message, for example in a telephone number. This is the fixed – wired systems technology.

We are mow experiencing the third phase. This has removed the inconvenienced condition of fixed location for the sending and receiving equipment. For example, a person interested in sending a message is within some very broad boundaries, free to choose where he will initiate the communication. In addition, there is no need to know the location of the person to whom he wishes to speak.

The development of mobile telephone thus incorporates various degrees of freedom on time and location in that one does not necessarily need to agree on absolute point in time but rather can to some degree negotiate over where and when to meet. Hence the quick switching over to mobile communications worldwide.

TODAY'S TELECOMMUNICATION NETWORK

Today's network can be grouped into two major parts:
Fixed-wired which is the old network that has been in place for some time(and the technology is fully understood to the extent that most of its limitations have been identified), and the wireless communication system side. Looking at the attatched diagram,the typical fixed-wired central office(CO) contain all the the intelligence and the

automatic systems that can handle calls, including the service- logic,call control, and call switching functions.
The logic systems and control were incorporated for easy upgradability; rather than upgrading each of the COs in the service areas which is more expensive and time consuming.

The fixed-wire system, which basically includes a network elements platform called the service control point (SCP) that handles the service logic, an intelligent peripheral (IP) provides specialized resources as text-to-speech, voice recognition and voice announcement, which are used during processing call.

The fixed-wired system with its distributed intelligence has many features. However, the biggest limitation of fixed-wired system is its limited support for mobility. Wireless networks began to develop because of the critical need for mobility. Advanced automation or intelligence had to be put into a central location in the network in order to provide mobile communication services that allows a customer based in Enugu in Eastern part of Nigeria to switch on his phone to get the package of services when in a different city like say Lagos or Minna in Nigeria.

Wireless networks are inherently rich in mobility management but generally poor in supporting features when compared to fixed-wired networks. In mobile communication systems all the features in wireless networks are provided by the Wireless Central Office also called the Mobile Switching Centre (MSC), which is potentially working in conjunction with the Visitor Location Register (CLR) and the Home Location Register (HLR).

From the diagram, one can observe that the wireless side and the fixed-wired have some common elements such as PSTN that carries
voice traffic between the two and the signaling system using Spread Spectrum technology (SS7). This carries the signaling messages between the elements.
However,the fixed-wired system uses IN application protocol(INAP) in processing a call, while the wireless uses cellular intersystem operations protocol IS41 for its Mobile Application Part (MAP) protocol
This arrangement enables the two networks to interoperate and able to communicate during a call set up[3].

CHALLENGES FACING THE WIRELESS INDUSTRY

Most of the expectations of the users and other expected innovations in the industry are seen as challenges by the industry. However, one of the biggest challenge facing the wireless industry is related to technology or technological innovation, that is how to achieve a seamless operations and synchronization regardless of whether the customers use wireless or fixed-wired network. The challenge is how to incorporate Intelligence Network (IN) technology of the fixed-wired system(e.g Service Contol Points) and the other intelligent peripherals into the wireless network in order to generate enhanced wireless services while still maintaing its mobility functionality[3] .

By the time they overcome this challenges, it will enable customers to receive services with the same look and feel in one seamless and smooth operation regardless of whether they use wireless or fixed-wired network. Similarly, it would enable wireless services providers to offer services with less effort and lower deployment costs.

Other challenges are also part of the customers' expectation which includes:
Wider coverage, higher mobility anywhere anytime; better quality of service in the areas of voice quality; noise control; noise reduction; blocked or dropped calls; low tariff.
I believe these challenges are not insurmountable. In fact the transition in mobile phones technology from 1G to 2G to 3G to 3G+ were as a result of limitations and challenges the telecommunication industry are continuously facing.The innovations are due to a need for a better services at higher data rates.

IMPORTANT TRENDS IN THE WIRELESS INDUSTRIES

Big challenges are drivers of change within a particular industry. There are some mega trends in the mobile industry today and these trends are having significant effect on the wireless communication business.

INCREASING DEMAND

The first wireless mega trend is in the area of customer demand, i.e. increasing customer demand. The acceptance of wireless service and phones has been phenomenal. The demand is far beyond expectation even in the developing nations. Communication has become inevitable for individuals, families, associates, and business organization worldwide. Therefore customers' expectations are increasing, as is demand for services.

STIFF COMPETITION

The second mega trend is keen competition. In every market around the world, wireless competition is increasing. All over the world including the emerging economies, there are stiff competitions among various mobile communication companies. A growing number of equipment suppliers are emerging from all the developed nations especially, USA, Europe and some parts of Asia to serve the ever expanding market. The intensifying competition has also driven the cost of services downwards to the benefit of the customers.[.]

DECREASING REGULATION / DE-REGULATION

Decreasing regulation is the third mega trend. In many nations of the world, de-regulation is a major factor in the growth and the development of telecommunication. Most of the monopolistic Telecommunication Acts are being repealed and replaced with more liberal ones that allow healthy competition within the industry. The de-regulations also opened-up the industry to good investors especially in the developing nations like Nigeria.

It was after the de-regulation that more Nigerians began to have access to Telecommunication facilities. The benefit of deregulation cannot be overemphasized. It also help to boost the economies of the Telecommunications industries worldwide. The world Trade Organization (WTO) trade agreement of 1997 [7]is a landmark agreement that breaks the barriers and many other countries follow suite.

DIGITIZATION

A fourth important trend is digitization. In reality, it was the choice of digital transmission that gave birth to wireless system technology . Because of continuous investment in digital technology research, new digital systems and digital transactions have been deployed rapidly in hundreds of markets for example the three basic.

second generation (2G) digital cell phone systems are in wide use today. Two of them use Time Division Multiplexing (TDM) and the third uses Spread Spectrum (SS). The TDM systems are the GSM and the IS-136 standard for TDMA (time division multiple Access). The SS system is code division multiple access. All these systems are based on digital technology. Various communication network circuit design configuration and interfacing have become possible after the development of Analogue–Digital Converter (ADC) and Digital – to – Analogue Converter (DAC) modules, components and systems.

INTERNET REVOLUTION

Internet revolution is the fifth mega trend in wireless communication. The Internet technology has created a global information infrastructure largely outside of government regulatory boundaries. It is not surprising that it has grown so quickly. Therefore one of the expectation from the mobile industry is to exploit the huge technological opportunities in the Internet and use it to bring commerce, entertainment and information research at the touch of a button.[4] .

These innovations will revolutionize the mobile industry as people come to expect access to those services from a mobile or remote location. The Internet is the fastest growing service, faster than cellular telephones or Personal Computers (PCs)[5]

There are other mega trends like minute migration, the broadband revolution, the information technology revolution,etc, diversifying subscribers' equipment for example using various configurations and equipment adapters.

CUSTOMERS EXPECTATIONS

There are many ways through which innovation can take place in any industry or product. It may be customers driven, it may be technological driven. This brings us to the issue of customers' expectation from the mobile industry. In fact, the future of any organization producing product or delivering services is tied to the organization meeting its customers expectations. Therefore, the future of mobile industry is tied to how effectively the industry can meet its customer's expectations. Most of the expectations of the customers of mobile or wireless industry are also challenges to the industry. Some of the expectations include:

1. Seamless operation regardless of whether the customer is using wireless or fixed wired network.
2. The industry should be aware that levels of service that were acceptable in 1980s are no longer acceptable to today's users they need to brace up for more stringent demand and be ready to deliver quality services.
3. Customers expect wider coverage and higher mobility anywhere anytime.
4. Customers expect higher speed of processing.
5. Customers expect longer battery life for the mobile phones and other hand-held devices.

6. Customer expects minimum energy consumption.
7. Customers expects clear video picture with distinct colour i.e. higher level matrix of pixels.
8. More friendly instruction or user manuals for the equipment and devices.
9. More features and functionality.
10. Faster call set up, privacy and ease of use.
11. Smart messaging and numerical paging.
12. Increased connection between people, computers and information. Not only people speaking to people, but people are speaking with machines; machines are also speaking with machines, and information data bases are showing information.
13. Lower tariff.
14. Reduced cost/bandwidth

For example, in the nearest future, I should be able to call my colleague in the office on his PBX phone using my mobile phone.
In medicine, there are many opportunities in wireless communication for innovation to enhance health care delivery. These are areas for development for the wireless industry.

TELECOMMUNICATIONS INDUSTRY STATUS IN NIGERIA

Telecommunication Industry development in Nigeria has been static (just like in some developing countries) until recently when the Federal Government decided to de-regulate the sector and this has enabled many Nigerians to enjoy Telecommunication facilities.

Before the era of de-regulation, Telecommunication Infrastructures were owed and controlled solely by the Federal Government. At that period, only about 15% of the entire population had access to telephone facilities[14] and the communication Networks were predominanatly fixed wired.

IMPACT OF DEREGULATION

Almost immediately after the de-regulation, investors started to invest in the telecommunication industry; within three years of De-regulation, several operating licenses were issued to interested investors and within a short period, few private communication companies have sprung up, both local and foreign and established Telecommunication networks and facilities in major cities in Nigeria. Even though there are few compared to the total population, it was much better than when the Industry was monopolised by the Federal Government.

MOBILE COMMUNICATION INDUSTRY IN NIGERIA

In many developing countries, established communication infrastructure were seriously, lacking, therefore the available telecommunication companies are flocking to wireless communication,thereby leap - frogging over fixed wired technologies. In Nigeria, among the privately owned Telecommunications companies, four of them are in mobile communication Business. They are: MTN; Globacom; Celtel and M-tel; all the four company have infrastructures to support 2G and 2.5G digital cellular phones. Just like the developed countries, wireless communication industry is the fastest growing business in Nigeria to day[15].

In Nigeria, there about 30 Million cellular phone subscribers, going by the statistics released by the mobile communication companies on their subscribers[13] .

This translate to about 20% of the entire population that now have access to telephone facilities within five years of their inception. This is a very big improvement when you compare it with almost fifty years of Telecommunications Industry in Nigeria and with only 15% of the population having access to the telephone facilities.

THE FUTURE OF MOBILE INDUSTRY IN NIGERIA

As I have said earlier, the future of wireless industry or mobile Industry in Nigeria is strongly tied to meeting the expectations of the customers. In Nigeria today almost every Telecommunications subscriber has switched over to wireless networks because of the opportunity of high mobility. There is continuous demand for communication service, the wireless Industry is growing at exponential rate. Therefore the following are the expectations of the users including the potential subscribers.

1. Better Quality Of Service.
2. Lower Tariff Is Also Expected
   In US, Europe and other parts of the world, as the subscribers increase in population, the cost of services drops proportionately[5]. But in Nigeria , it has not been so, despite the fact that most of the mobile communication companies have broken-even. Their tariffs are still very very high. The future of the industry depends on how well they can meet all these expectations.
3. Seamless Connectivity And Better Interoperability Among The Wireless Networks.
   It is not uncommon to see some Nigerians in big cities carrying two or three mobile phones at the same time. This is because inter - connectivity between different networks is still very poor for example if a customer is on MTN network and he wants to call on a person who is on another network different from MTN, the call may take one minute or two minutes to go through. These are areas of challenges to the Industry and they must be addressed as soon as possible for the sake of the future of the Industry.

4. Expectations In The Areas Of Research And Development
   In the area of research and development, the mobile Industry in Nigeria is not doing enough; it is expected that the mobile industry should take keen interests in research and development. The industry should actively support and encourage research in to new areas in wireless communication networks e.g Intelligient Networks(IN); Satelite Telephony Technology and a host of others. This can be done by collaborating with some Universities and polytechnics to engage in progressive and productive research directed at addressing the challenges facing the industry especially in the energy and power supply system and other essential infrastructures that are vital to the growth of Telecommunications in Nigeria.

5. Expectations In The Areas Of Training And Technology Transfer.

Because the future of the Industry depends on the qualities of their staff: engineers, technicians,managers, administators etc.For example,the London Sunday Times reported how the founder of the ‘Source’,Daniel Mitchell developed and grew his business from zero to sales of 35Millon Pound. Mitchell argues that “success is about customers, but it is also about the people you employ” (London Sunday Times,2004)[11]..

It is expected that the Industry should invest heavily on staff development programmes and look in-wards for the recruitment of young and dynamic ones. The industry should encourage students in the higher institution by taking them for industrial attachment from time to time. In Nigeria, the mobile communications have not been doing this.

6. Other Expectations

These include the convergence between the wireless networks and the internet technology innovation to enhance the quality of their services and also to bring down the cost of their services.

Customers and the society at large expect that the government should urgently address the issue of Energy and power system. This will allow the mobile industry to conserve some resources thereby bringing down their tariff.

CHALLENGES FACING WIRELESS INDUSTRY IN NIGERIA

The biggest challenge facing the mobile industry in Nigeria is Energy especially issue of electric power supply. As at today, the companies within the industry generate their own power through standby generators located in each of their base stations.

This is considered outrageous and makes the whole operations and services very expensive.

Other challenges include poor telecommunication infrastructure and inadequate skilled man power. Most of the customers expectations are also challenges to the industry.

CONCLUSION.

Writing about the future of Mobile Industry, this paper had attempted to discuss the reasons while mobile communications was widely accepted.The paper identify some of the emerging technologies that were developed in response to the demand of wireless communication

The paper also touched on some of the challenges facing the engineers and professional in the industry and how these challenges could be combated through technological innovation coupled with extensive research and development.

In the paper , we identified some very challenging and important trends in wireless communication and how are affecting the industry…..

In conclusion, we believe that

the future of wireless industry world wide will deliver better and quality services to its numerous customers through keen competition among the the telecommunications companies. The future will also witness a lot of innovation through dedicated research in advanced automations and Intelligent Networks in other to improve their services and also to make the service more affordable to its Millions of potential customers especially in the developing Nations of the world

Future of Mobile Industry will be technology and market driven such that there will be convergence of internet engineering, wireless communication networks and Mobile Computing..

The future is also very promising economically as the customers will continue to demand for communication services world wide especially wireless communication services..

However among the numerous wireless communications companies, it is those that could invest in productive research leading to innovations that will respond to the expectations of their customers that will eventually survive the imminent competitions.

REFERENCES:

1. Edwin Wright, Deon Reynderes, Practical Telecommunications and Wireless Communication for Business and Industry, 1st Edition,2004, Newness, UK.
2. Louis E. Frentes, Principle of Electronics Communication Systems, 2nd Edition, Mc Graw Hill
3. Ramachandria P. B., Wireless Intelligent Numerous: Gateway to the Growth of Enhanced Wireless Services, 1998 Annual review of communication
4. F. Graig Fasrwill, MEGATrends in the wireless Industry, 1998, Annual review of communication
5. Hart E.T.,The Wireless Stampede,IEE/FSA International Semiconductor Executive Forum, May 2005, pg 149
6. Major Mobile Phone System, 2002, Available [http://www.radio-electronics.com]
7. WTO's Landmark Agreement n Basic Telecommunications Services, ITU News Nov. 4, 1997, pg 3438
8. The International Engineering Consortium (IEC), Global System for Mobile Communication, 2000, Available [http://www.iec.org]
9. Eric Senunas, The 1997 GATS Agreement on Basic Telecommunications, A Triumph of Multilateralism or the Market? Nov. 4, 1997
10. San Diego Regional Technology Alliance, Wireless Wonders, Accessing San Diego Wireless Industry, 2000
11. Paul Trott, Innovation Management and New Product Develpoment,3rd Edition,2005, Prentice Hall.,
12. Trends in Telecommunications Market in Nigeria, 2003-2004
    Editorial By:Corporate Planning & Research Team, NCC, Abuja
    Available [http://www.ncc.gov.ng]
13. Mr. Larry Chidi Maha,Internet and Mobile Communications,2004
    Available [http://www.ncc.gov.ng]
14. Eng. Ndukwe, An Overview of the Nigerian Telecommuncations Envirronment,ITU Telecomms 2004, available[http://www.ncc.gov.ng]
15. Steve Wallage, Nigeria Much More Growth to come,May 2005,available [http://www.ncc.gov.ng]

3G (GSM) TERMINAL CONCEPTS FOR RURAL AND URBAN EMPOWERMENT

O. E. AGBOJE

Department of Electrical and Information Engineering, Covenant University
oeagboje@yahoo.com

Abstract

The author is of the opinion that the introduced 2G GSM services were designed to be consumable services and not wealth generating ones. This paper is set to show how third generation (3G) GSM (hereafter 3G) with a well designed terminal technology can be used to transform the population from information consumers to information producers and users, which would lead to economic and social transformation of the population. The purpose of this paper is to look at some terminal design concepts, which could make 3G a development resource for community empowerment in Nigeria.

1. Background
1.1 Narrowband Telephony
The original concept which governed the introduction of GSM in early 1990s in Europe was that of narrowband subscriber telephone network [1, 2, 3]. The services provided can be described as consumable and not capital resources. At that time, the wideband services which are presently available under 3G were just dreams to be realised in very distant future [4, 5].
The introduction of 2G GSM into Nigeria in 2002 was a tremendous success. At this time of writing, the 3 dominant GSM networks, MTN, Celtel and Globacom continue to spread into remote villages and communities. Nevertheless, this widespread has not generated much needed economic benefit to the population.

1.2 The 2.5G Era in Nigeria
This complex technology is yet to reach the average user. Unlike 2G, certain 2.5G services enjoy content-based billing. Examples of 2.5G content-based billing services include: GPRS (General Packet Radio Service) and EDGE (Enhanced Data for GSM Evolution) [1, 5].

1.3 The 3G Technology as an Economic Driving Force
A 3G billing method is expected to give it a wider acceptance than 2.5G GSM in Nigeria [6,7,8,9,10]. A country with a failed PSTN (Public Switched Telephone Network) such as Nigeria is a potentially good market for 3G technology. The country is seriously in need of widely available broadband terminal services such as: Internet browsing, Internet multimedia services, Internet video conferencing, etc, which are offered by 3G technology. The 3G services are expected to change many subscribers from information consumers to information producers at both urban and rural levels of the country. It would also mean information empowerment, which would lead to economic empowerment and job creation.
As Internet gateway, 3G terminal services will compete reasonably with the present conventional Internet Service Providers (ISPs).
For this to happen, appropriate 3G terminals must be put in place. This is expected to spread the technology to most remote parts of the country.

The importance of an appropriate and well designed terminal is made abundantly clear when millions of US dollars was recently invested in the development of 3G terminals in China [11, 12].

2. Some 3G Services for Rural and Urban Community Empowerment.

The economic empowerment inherent in 3G services must not be allowed to waste without relevant and well structured approach. The services of interest in this paper are mainly those designed to create gainful employment and economic development of both urban and rural parts of Nigeria.

2.1 Web-aided Market Sourcing Services

Unlike 2G GSM, Internet services will be widely available under 3G. For the rural communities, a direct Internet link between them and their Internet market partners will be provided. Through this, greedy middlemen can be omitted from business deals.

2.2 Internet–based Direct Procurement Services

This service will link urban and rural communities with producers of capital goods such as industrial plants for direct procurement at much reduced prices.

2.3 Internet–based Distant Learning Services.

For the rural communities, access to Open Universities and other relevant institutions for lecture download and online examination services through 3G terminal can be described as academic empowerment of the best kind.

2.4 Internet–based Distant Library Services.

One of the most relevant tools for rural and urban population empowerment is learning resource such as library services. However, these services are hardly available where they are most needed in Nigeria. This is where 3G Internet services can be of crucial importance. With such a system in place, students from rural communities can access and make use of distant library services world wide and empower themselves for a better tomorrow.

2.5 Internet–aided Telephone and Fax Services

Certain voice on Internet protocol (VoIP) telephone service providers such as Net2phone, have over the years perfected Internet low cost telephone services for Internet users in metropolitan areas.

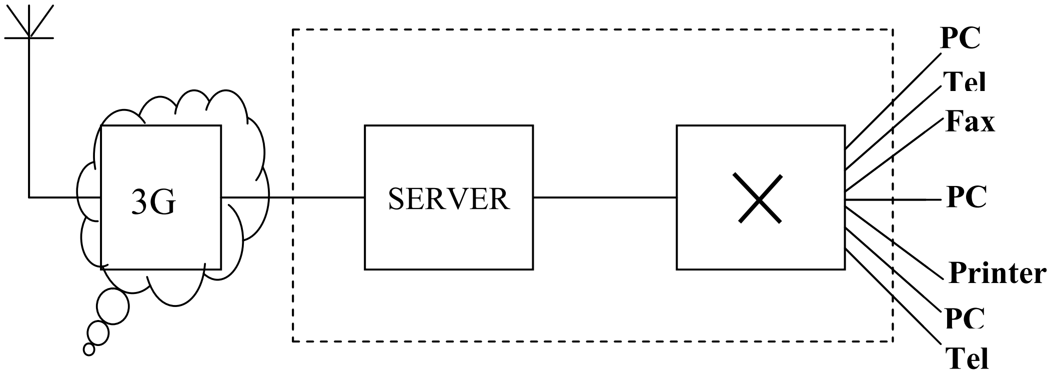
With the introduction of 3G in Nigeria, the rural population will also have access to these services but at reduced cost [13, 14]. Terminal concepts for these services are described in sections 3.3.

2.6 Direct Community Development Services

With these services, rural communities can source for information concerning international technical aid bodies and directly seek for help when needed. This direct approach is likely to reduce response time of international organisations and Non Governmental Organisations (NGOs).

3. The 3G Terminal Concepts

Fig. 1.0: A 3G GSM Terminal



Over the years, several technical support networks and tools such as communication satellites, the Internet, personal computers, Global Positioning Systems (GPS), 2G GSM, etc, were introduced for human and technological development. Some countries reacted and designed appropriate interfaces or terminals to connect into these vital resources. An earlier example of a country's response to a global technological invention is the German T-Online. T-Online is an Internet software terminal, which greatly simplified the use of the Internet in early 1990s in Germany [15]. Generally, those that reacted appropriately to great inventions reaped maximum technological, economic and social benefits. Those that failed to respond remained left out. Response to major technical inventions should be a national policy. Failure to do this would mean an endless race of "catching up".

The introduction of 2G GSM in Nigeria was not followed by a national Terminal design policy. As a result, GSM could not become an economic driving force and capital good. The ability and capacity of a country to respond to global innovation such as 3G GSM is determined by the country's well developed technological response strategy, with which to reap maximum benefit from such global innovations. We must not repeat the mistakes of 2G/2.5G GSM.

3.1 A Simple 3G Terminal

A 3G network is a dynamic, flexible and yet very powerful technological system. If the present national coverage of GSM networks, which extends into very remote villages, can be repeated in the spread of 3G, simple laptops can be configured to act as terminal LAN network servers, which can be used by rural communities to access the internet, make phone calls, and directly sell their products in the global market.

At a 3G download data rate of over 2Mbps, several information files and multimedia content can be received and stored in the laptop memory for future use. The laptop approach is very appropriate for an environment with unreliable power supply.

Thus, with a simple laptop server, a printer and few VoIP (Voice over Internet Protocol) telephones, an information service provider with some support personnel can serve a rural community offering latest Internet services and linking the people to the outside world. And without leaving their villages, job seekers can search, apply and attend online interviews.

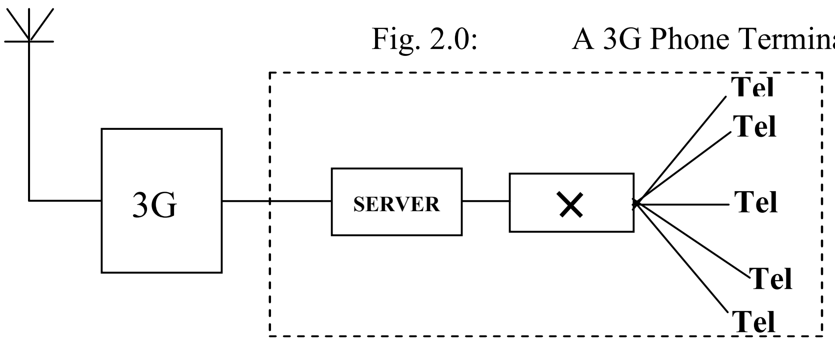
3.2 A High Capacity 3G Terminal

With 3G, a well planned terminal should enable a professional Internet service provider operates in large part of the country including many rural areas. This advanced 3G terminal is expected to offer Internet services, which are close or better than those presently offered by conventional ISP's. This is technically possible because 3G can provide data rates in excess of what obtains in some of the ISPs in Nigeria today. But as stated earlier, providers of 3G services must agree to give the necessary support to make 3G a useful resource for the development of Nigeria.

To offer reliable Internet services to rural and urban communities, the terminal must have the structure of an Internet LAN (Local Area Network). This terminal concept requires reliable and high speed servers, with which some of the best 3G services can be utilised. Services are expected to be delivered at a much faster rate than some of the ISP's. At the heart of the advanced 3G terminal concept is a high speed server. This concept can be judged as the most appropriate for an advanced network such as 3G.

3.3 Special Service Terminals

Fig. 2.0: A 3G Phone Terminal



For a rural community empowerment, offices and individuals must be offered low cost Internet telephone and fax services [13,14]. For this group, the terminal network described earlier must be extended to include phone and fax services from world class providers such as Net2phone.

Net2phone and similar low cost VoIP [4] telephone service providers are known for reliable phone and Fax services. To maximize these services, multi-user versions of the applications are required to be installed in the server followed by several telephone boxes. The telephone boxes are designed to give some degree of privacy to users. With this arrangement, users will be empowered to use phone and fax to run their business at relatively reduced cost. And like conventional cyber café, far more jobs will be created at both rural and urban areas than is presently possible under normal GSM [16].

3.4 Recommendation

To have a significant impact on the economy and empower the rural and urban communities of Nigeria, I strongly recommend an advanced 3G terminal as described in sections 3.2. As mentioned in the said section, such a terminal must offer good internet services comparable to those offered by conventional ISPs in metropolitan cities.

4. Conclusion

In this paper, the importance of appropriate response to world technological innovations, as it applies to 3G, was highlighted.

To make the introduction of 3G GSM bring much needed technological advancement and community empowerment, the paper strongly suggest well designed terminal networks for both rural and urban applications. It was clearly put that only an appropriate terminal design can reap maximum benefit from big inventions such as 3G GSM. The terminal concepts suggested can be realised with modern and available technologies.

The Federal Government of Nigeria and Stakeholders should seriously look into the issue of appropriate terminal design as recommended in this paper. Only the Government and licensing authorities can make licence applicants and telecommunication operators offer relevant and job creating services to the Nigerian people at an affordable price. Only Government and Stakeholders can make operator provide a good interface and protocol, through which a well designed terminal can access their network and empower the communities.

It should be noted that a deep sea port without excellent facilities for ships is a useless sea port. Ships will avoid such port and move elsewhere. Nigeria will rise up to new technological innovations and will not lose its place in a world governed by Internet and Telecommunication advancement.

Reference:

[1] Louis Frenzel E., “Principle of Electronic Communication System” Published by McGraw-Hill companies, Inc, (2003)

[2] Lawrence H., Richard L. and Geoff L, “GSM Superphones” Published by McGraw-Hill companies, Inc, (1999)

[3] Regis J., Bud B. and Donald W. Gregory “Voice & Data Communications Handbook, Fourth Edition” Published by McGraw-Hill companies, Inc, (2002)

[4] James Henry G., “The Irwin handbook of Telecommunications” Published by McGraw-Hill companies, Inc, (2000)

[5] William W., “The Complete Wireless Communication Professional: A Guide for Engineers and Managers” Published by Artech House, Inc., Boston London, (1999)

[6] Glo Mobile Nigeria ready to launch 3G at: http://www.mobileafrica.net/n1658.htm, 2006

[7] 3G Technology – Benefits to the Nigerian Economy at: www.ncc.gov.ng/Workshop%20Papers/3G%20Spectrum/Moto-NCC-24April.pdf, 2006

[8] Nigeria Announces Plans for 3G Licenses at: http://www.cellular-news.com/story/21345.php, Jan, 2007

[9] NCC To License Fifth GSM, 3G Networks at: http://www.independentngonline.com/news/47/ARTICLE/17343/2006-12-20.html, Dec., 2006

[10] Nigeria Communication Commission Consults Before a Final Push for 3G at: http://www.itu.int/ituweblogs/treg/For+3G+Nigeria+Communications+Commission+Consults+Before+A+Final+Push.aspx, April, 2006

[11] Vimicro, BIRD Collaborate to Develop 3G Multimedia Terminals at: http://www.3g.co.uk/PR/June2006/3179.htm, June, 2006

[12] Starent Brings 3G to China at: http://boston.internet.com/news/article.php/980281, February 25, 2002

[13] Net2Phone launches wireless VoIP handset at: http://www.robertlett.com/newspro/news.shtml, 2002

[14] Net2Phone Selects BridgePort Networks To Provide MobileVoIP Convergence Solutions MobileVoIP Convergence Enables Service Providers with a Wireless Solution for Consumers and Enterprise Users at: http://web.net2phone.com/about/press/releases/20060509.asp, May, 2006

[15] T-online at: http://www.T-online.de, 1990

[16] How Mobile Operators Can Monetize 3G Investments Through an Effective Applications Platform at: http://whitepapers.techrepublic.com.com/search.aspx?tags=mobile+operator, Jan 2006.