Covenant Journal of Informatics & Communication Technology

Vol.2 No. 1, June, 2014.

A Bi-annual Publication of the Departments of Computer Information Science, and Electrical & Information Engineering, Covenant University, Canaan Land, Km 10, Idiroko Road, Ota, Ogun State, Nigeria.

Website: http//Journal.covenantuniversity.edu.ng

It is a condition of publication in this journal that manuscripts have not been published or submitted for publication and will not be submitted or published elsewhere.

ISSN: 2354 – 3566

Published by Covenant University Journals,
Covenant University, Canaanland, Km 10, Idiroko Road,
P.M.B. 1023, Ota, Ogun State, Nigeria.

Printed by Covenant University Press.

Articles

AVAILABILITY, ACCESSIBILITY AND FREQUENCY OF USE OF ICT TOOLS BY HEALTH PROFESSIONALS IN ILORIN METROPOLIS

O. T. Afolayan
&
R. A. Oyekunle

Department of Information and Communication Science
Faculty of Communication and Information Sciences
University of Ilorin, Ilorin
ummusumayya1@yahoo.com

Abstract: This research explores the availability, accessibility and the extent of ICT usage by health professionals in Ilorin metropolis. Data was collected using a sample size of 174 health service providers in 34 public and private hospitals by means of a structured questionnaire. Findings from this study indicate that the ICTs that were commonly available and used by the health professionals are the older ICTs. These include TV set, DSTV, photocopiers, fax machines, and convergent technologies like mobile phones and computers. Although, Internet driven facilities like video-conferencing, teleconferencing, web discussion forums and email among others, were not readily available to the health professionals, a comparative analysis reveals that there was superior availability and use of computers, projectors, e-mails and Internet by health professionals in private hospitals than those in public hospitals. This is as a result of the financial implications of procurement and installation of these facilities. Interestingly, the level of awareness of users on the numerous benefits of ICTs on their job and productivity is quite impressive. It is therefore, concluded that hospitals' regulatory agencies should collaborate with International agencies such as World Health Organization (WHO), United Nations Organization (UNO) and the European Union, among others, to aid in the successful implementation and funding for the procurement of sophisticated ICTs to facilitate the dissemination of up to date health information to public and private hospitals.

Keywords: ICTs, Health professionals, Usage, E-mail, Teleconferencing, Video-conferencing

1.0 Introduction

The health sector is one of the largest information consuming sectors. As a result, access to up- to-date and timely information by health professionals remains a sine qua non for proper diagnosis, prevention and treatment of diseases. Readily available information that affects the well being of patients is always critical and over the last decade, the advent of Information and Communication Technologies (ICTs) have contributed immensely to the continuous learning, sharing and dissemination of health information amongst professionals. The United

Nations, in a report, corroborated the fact that ICTs have the capacities of contributing to health education, knowledge sharing, health monitoring, health statistics gathering and achieving the millennium developmental goals (UNDP, 2003). This encompasses the full range of ICTs from radio, television, communication resources such as fax machine, telephone (fixed and mobile), email, Internet, Personal Digital Assistant, manual and computerized data systems, databases, satellite technology and so on. It is therefore essential for health professionals to have access to ICTs and use them to update their knowledge and skills in order to be relevant in the medical profession and also to conform to best the practices.

Some studies have shown that despite the immense benefits of ICTs as a means of delivering quality health care services, the potential of ICTs have not been fully harnessed by health professionals especially in developing countries. This is due to problems of access (slow or unreliable Internet connectivity), high subscription cost of information materials; lack of awareness of what is available; lack of relevance of available information (i.e. not meeting professionals' needs in terms of scope, style, or format); lack of time and incentives to access information; and lack of interpretation skills (Gatero, 2011). These enumerated problems may affect the availability, accessibility and use of ICTs by health professionals particularly in developing countries. Furthermore, Idowu (2003) stated that most of the health institutions in developing countries are battling with some identified barriers such as high cost of ICT equipment, power failures, and inadequate telecommunication facilities. Except these challenges are promptly attended to, the benefits of ICTs may not be realizable both at the short and long run. In essence, the myriad of problems in the health sector may not be totally eliminated by merely acquiring these technologies, but harnessing the technologies for development.

This study therefore, intends to explore the available ICTs, accessibility and extent of use by health professionals in public and private hospitals in Ilorin Metropolis. Such benefits which include making better treatment decisions; providing higher quality and safer care; national and local health information systems supporting the development of effective, efficient and equitable health systems, among others can achieve the desired outcomes only if ICTs are available, accessible and used accordingly (Davies et. al. 2006).

1.1 Statement of Problem

Universal access to information by health professionals is a prerequisite

for meeting the Millennium Development Goals and achieving health for all (Godlee et al, 2004). Health related millennium goals center on improving the health of the poor and the marginalized, combating child mortality rate and the prevention of diseases such as HIV/AIDS. On this note, many health institutions in developing nations are now harnessing the potentials of ICTs to improve the delivery of quality health care services. However, it has been observed that despite the high cost of ICT equipment, many health institutions are still struggling and investing heavily on procuring the tools considering their potentials. This high cost is limiting the availability of the tools and consequently affecting access to the tools and frequency of use by the health professionals. Based on this premise, the research questions can be stated thus:

1. Are ICT tools available for use by health professionals in Ilorin metropolis?
2. Are these tools accessed at variance by health professionals in Ilorin metropolis?
3. What is the extent of use of ICT tools by health professionals in Ilorin metropolis?
4. What activities do the health professionals in Ilorin metropolis undertake with ICT tools?
5. How do health professionals in Ilorin metropolis perceive the impact of ICT tools on their job functions?
6. What are the likely constraints in the utilization of ICT tools by health professionals in Ilorin metropolis?

In view of this, the overall objective of this research is to determine the availability of ICT tools, level of accessibility and extent of use by health professionals in Ilorin metropolis using sampled public and private hospitals. Specifically, this research paper will address the following sub- objectives:

1. Determine the state of ICT facilities in selected public and private hospitals in Ilorin metropolis.
2. Find out the activities ICTs are used for by health professionals in Ilorin metropolis.
3. Determine the places of access to ICTs by health professionals in Ilorin metropolis.
4. Determine the impact of ICTs on the job functions of health professionals in Ilorin metropolis.
5. Evaluate the attitude of health professionals in Ilorin metropolis towards the use of ICTs.
6. Determine the constraints to the utilization of ICT tools

by health professionals in Ilorin metropolis.

7. Suggest policies and strategies that will support ICTs interventions in health institutions in Ilorin metropolis.

2.0 Literature Review

2.1 Successful ICT Interventions in the Health Sector

The application of information and communication technologies tools (ICTs) or alternatively e-health tools to successful health care delivery had been widely demonstrated through various ICT health projects in developing and developed world. E-health, an umbrella term that encompasses the use of ICT in healthcare, includes telemedicine, where medical advice or consultation is provided over long distances via Internet, radio, telephone and other communication devices. In another vein, e-health is the use of ICT tools for dissemination of health related information such as HIV/AIDS and vaccination hubs using radio, television, Internet or short message service (UNDP, 2007) Successful application of e-health tools in the developed economies had contributed to the delivery of quality healthcare services. For instance, rapid access to health information can easily be achieved via the World Wide Web (WWW), Internet and other relevant facilities like online databases. On the contrary, the utilization of e-health tools in the developing countries are posed with challenges which needs to be addressed. These barriers include lack of computer equipment, lack of computer skills, cost of computer and so on. These have contributed to the unreliability, inaccessibility and lack of sustainability of ICT infrastructures (Ouma and Herseman, 2008).

E-health projects that were supported by organizations such as World Health Organization (WHO), United Nations (UN) European Union (EU) and various Non-Governmental Organizations (NGO's) have led to the successful implementation of e-health solutions especially in the developed world. For instance, EU supported projects had placed Europe in a world leading position in the use of health networks, Electronic Health Records (EHR) and deployment of health information systems.

Specifically, ICTs had contributed quantitatively to the health sector in some developing economies through ICT pilot projects. For instance, Greenberg (2005) indicated that a number of ICT pilot projects have demonstrated improvement such as 50% reduction in mortality or 25-50% increase in productivity within the healthcare system. E-health projects had been successfully implemented in countries like Indonesia, Philippines and Thailand. In Indonesia, an integrated web based GIS was designed to monitor and detect outbreak of dengue fever amongst children. Similarly, in the

Philippines, a Community Health Information system was designed to disseminate health information to the community and public at large. It is worthy to note that these individual projects were implemented in phases and improved on overtime. In Thailand, a low cost hospital information system was designed for rural hospitals to perform certain functions such as work flow management, patient appointment, billing and patient registration. In order to curtail cost, free open source software applications (FOSS) were utilized in a networked environment linking fifty (50) rural hospitals to facilitate knowledge sharing (UNDP, 2007).

Similarly, in Uganda the effective use of ICT had prevented avoidable maternal death. In Bangladesh and India the global satellite technology combated the outbreak of epidemics and ensured the effective prevention and treatment through adequate online health information. In South Africa, three health information systems were developed. The first was a National Health Care Management Information System (NHC/MIS). This system was designed to track medical records, patient registration, billing and scheduling in some selected hospitals. The following information was generated from the system: patient record, ID verification blood group, last ten (10) diagnoses, treatment prescription and medical aid (Electronic Government, 2004). Secondly, an information system was also implemented in the South African Province of Limpopo which consisted of 42 Hospitals (2 mental health facilities, 8 regional facilities and 32 district facilities).The information system was designed to improve patient care, manage efficiency of hospitals and increase the quality of services, track patients records, admission, discharges and transfers, appointment etc. This project was successful only in few of the hospitals; some factors identified for the failure include not taking into account the social/professional culture of health Organizations, lack of education of users, failure of developers to identify and learn from past projects. Thirdly, the South African District Health information System (DHIS) is a systematic data gathering tool that was used to identify public health issues. The Health information system programme (HISP) model was adopted in implementing DHIS. The implementation steps were categorized into six steps: Step 1- Establishment of district into teams, Step 2-Performance of an information audit of existing processes, Step3- Formulation of operational goals, Step 4- development of system and structures to support data handling, Step 5- Capacity building of health care providers. Step 6- Development of an information culture. The HISP model was successfully

implemented in all the districts (Electronic Government, 2004).

In view of the above, Davies et al. (2006) argues that the successful implementation of health information system must consider the following factors: pay attention to past experiences, involve users in the planning and design of the system, build information culture, strengthen capacity of users, set realistic goals and lastly focus on the benefits of the system rather than the technology. In concrete terms, the utilization of e-health tools such as health information systems, websites, electronic health records, medical databases by health professionals will facilitate the following benefits identified by Davies, Trude, Mcconnell, Ramirez, Shields, Drury, Kumekawa, Louw, Fereday, and Nyamai-Kisia, (2006):

1. Improve dissemination of public health information;
2. Enable remote consultation, diagnosis and treatment through telemedicine;
3. Enhance collaboration and cooperation among health professionals by sharing of learning and training resources;
4. Support more effective health research and its dissemination and access to research findings;
5. Strengthen the ability to monitor incidence of public health threats and respond in a more timely and effective manner;
6. Improve the efficiency of health systems.

2.2 Related Works

Quite a number of studies in the literature have evaluated the use and impact of ICT on the health sector in developing countries. Some studies that have considered the use of ICT's in Nigerian Teaching Hospitals include (Idowu et al 2003; Adeyemi & Ayegboyin, 2004; Olatokun & Adeboyejo, 2009) among others. Specifically, Idowu et al (2003) identified ICT indicators such as mobile phones, personal computers, non-hospital Internet facilities in four Nigerian teaching hospitals: Obafemi Awolowo University Teaching Hospital Ile-Ife, Ladoke Akintola University Teaching Hospital Ogbomoso, Lagos State University Teaching Hospital and University College Hospital Ibadan. Amongst the ICTs that were readily available, mobile phones were highly utilized by almost all the medical experts in the teaching hospital; they were essentially used for communication consulting and generally for improving health care services. In the surveyed hospitals, there was lack of Internet connectivity which greatly impaired health delivery services. On their own part, Adeyemi and Ayegboyin (2004), in a survey involving four general hospitals, 10 primary health-care centers, and six private hospitals in Nigeria reported that none of the

institutions had e-mail access or a Web site and only very few of the workers possessed personal computers or had any measurable computer skills. On the contrary, Olatokun & Adeboyejo (2009) reported a 100% Internet usage by (Nurses and Pharmacist), 98.1 % (Medical doctors) and 96.2% (Nurses/ Midwives) at the University College Hospital (UCH) Ibadan. This clearly shows that Internet facilities are becoming entrenched and indispensable in the hospital environment.

Similarly, Ouma & Herselman (2008) investigated ICT infrastructures and e-health technologies in place in five rural hospitals in Kenya. A technology assessment was carried out within the five rural hospitals focusing on the access level of ICT, current condition of ICT infrastructure and the barriers militating against the use of ICT. The findings from their study revealed that ICT infrastructures that existed include computers, Internet, and information systems in the pharmacy and finance offices respectively. However, Electronic Health Records (EHR) and tele-medicine were not in place. Barriers to the successful implementation of ICT in these rural hospitals include lack of computer equipment, lack of computer skills and cost of computer. These barriers affected the reliability, accessibility and sustainability of ICTs in the rural hospitals. It is note worthy to realize that factors such as attitude problem, cultural/environmental factors are likely barriers that can militate against the successful implementation of ICTs in a rural environment.

Another study by Adedoyin, Imam and Oladapo (2009) investigated the ICT literacy among the health workers of Igbinedion University Teaching Hospital (IUTH) Okada, using a survey method to elicit information from the respondents for the study. Among other things, the study revealed that a greater number of respondents which represents 74.5% of the sampled population are aware of the use of ICT equipment in health institutions; though media and friends were the major sources of awareness to the respondents. The study also revealed that 66% of the respondents were aware of tele-medicine but unfortunately 76.6% do not have ICT equipment in their various departments. The study also investigated the attitude of health workers in IUTH to the use of ICT and it was revealed that non-availability of ICT equipment had hindered the health workers from showing serious interest, leading to a widespread apathy towards ICT because of lack of facilities for their use. The study concluded from its findings that a greater number of health workers in the developing nations are becoming aware of the use of ICT in health institutions to

increase the effectiveness of health care services delivery.

Jimoh, Pate, Lin, and Schulman (2012) also investigated the potential of information and communication technology (ICT) adoption among maternal and child health workers in rural Nigeria. A prospective, quantitative survey design was used to collect data from quasi-randomly selected clusters of 25 rural health facilities in 5 of the 36 states in Nigeria over a 2-month period from June to July 2010. A total of 200 maternal and child health workers were included in the survey, and the data were analyzed using a modified theory of acceptance model (TAM). The results indicated that there was no significant difference between ICT knowledge and attitude scores across the states. There were significant differences in perceived ease of use (P<.001) and perceived usefulness scores (P=.001) across states. Midwives reported higher scores on all the constructs but a lower score on endemic barriers (which is a more positive outcome). However, the differences were only statistically significant for perceived usefulness (P=.05) and endemic barriers (P<.001). Regression analysis revealed that there was no interaction between worker group and age. Older workers were likely to have lower scores on knowledge and attitude but higher scores on perceived ease of use and perceived usefulness. It was also revealed that worker preference for ICT application in health varied across worker groups and conflicted with government/employer priorities.

In contrast to what obtains in public and private hospitals in developing countries, the basic ICT infrastructure consisting of computers and Internet connections is today available in most of the General Practitioner practices in Europe (Empirica, 2008). The electronic storage of administrative and medical patient data, the use of computer during consultation with patients and other uses of ICT in the health area are becoming more and more a daily experience among General Practitioners in Europe. However, there is still room for improvement when it comes to electronic networks connecting their IT systems with other health actors, the electronic exchange of patient data and electronic interactions with patients. Today, almost all General Practitioner (GP) practices in the European Union use a computer, thereby benefiting from all eHealth has to offer.

3.0 Methodology

Hospitals under the National Health Insurance Scheme (NHIS) were used as the population of study from which 50% of the hospitals in Kwara central senatorial district were randomly sampled. NHIS was established under Act 35 of 1999 by the Federal Government of Nigeria to register licensed government or private health care practitioners or facilities, for the provision of easy

and affordable healthcare to all Nigerians. The instrument for data collection was mainly questionnaire developed and designed based on a 22 item criteria and divided into 7 sections. It was examined to ensure content validity and reliability within the target context. A pre-test of the questionnaire was performed using 3 experts in the ICT area to assess logical consistencies, ease of understanding, question item sequence, adequacy, and context fitness. Necessary modifications were made according to the comments collected from these experts. Furthermore, the reliability of data collected was validated using Cronbach's alpha statistics for each of the sub scales, and the results is as follow:

Table 1: Reliability Results

Construct	No. of Items	Cronbach's Alfa
ICT Facility Used and Extent of Use	2	0.812
Task and activities carried out with ICT	2	0.836
Access to ICTs	6	0.892
Impact of ICTs	1	0.676
Attitude Towards ICTs	2	0.852
Constraints to the Use of ICTs	3	0.809

The first section of the questionnaire consists of demographic profile of respondents. The second section of the questionnaire was on ICT facilities used and extent of use, the third section consists of questions on tasks and activities carried out by health professionals using ICTs, the fourth section was on access to ICTs, the fifth section contained questions on impact of ICTs on the job functions of health professionals, the sixth section was on attitudes towards ICTs and the seventh section was on constraints to the use of ICTs. The questionnaire were randomly distributed within the research population, a total of 250 questionnaire were randomly administered within three months, 210 were received, and 36 of them were not useable. The remaining 174 valid and complete questionnaires were used for the quantitative analysis. These represented a useable response rate of 69.6%. The data collected was analyzed using the statistical package for social sciences (SPSS) version 15.0.

4.0 Presentation of Results

Table 2 below shows the demographic information of the health professionals. Female health professionals were more than twice their male counterparts; this is as a result of the fact that females were more in the nursing profession than males as revealed by a cross tabulation analysis of gender and category of health professionals. A high percentage of the health professionals are within the age of

26 and 35 years and about three quarter of the respondents had between 0-10 years experience, this implies that there were younger health professionals in the study area. Also, there were more health professionals in private hospitals than in the public hospitals signifying that private hospitals are more in number in the metropolis. Nurses were more in number as naturally expected and the few, neither doctors nor nurses, were laboratory technologists, pharmacists, and community health workers. All the health professionals have the minimum qualification for their respective professions with only 2 of them having additional qualification in other areas i.e. FRCOG (Fellowship of the Royal College of Obstetricians and Gynecologists).

Table 2: Demographic Information of Respondents

Variables	Frequency	Percent
Gender		
Male	54	31.0
Female	120	69.0
Age		
26 – 30yrs	63	36.2
31 – 35yrs	63	36.2
36 – 40yrs	19	10.9
41 – 45yrs	9	5.2
46 – 50yrs	14	8.0
51 and above	6	3.5
Type of Health Institution		
Public	73	42.0
Private	101	58.0
Category of Health worker		
Doctor	57	32.8
Nurse	97	55.7
Others	20	11.5
Highest Educational Qualification		
Registered Nurse	47	27.0
Registered Nurse Midwife	36	20.7
B. Sc Nursing	14	8.0
M. Sc Nursing	1	0.6
PhD Nursing	1	0.6
MBBS	53	30.5
Additional degrees in other areas	1	0.6
Others	21	12. 1
Years of professional experience		
0-5	85	48.9
6-10	45	25.9
11–15	18	10.3
16-20	3	1.7

21-25	11	6.3
26–30	3	1.7
31 and above	4	2.3

ICT Facility Used and Extent of Use

The result of the analysis revealed that the ICT facilities and ICT enabled services that had more than 40% availability and functionality include DSTV (66.1%), TV set (87.4%), mobile/land phone (82.8%), computers (75.9%), printers (62.6%), email (50.6%), Internet searches (43.1%). Others like fax machine, video camera, multimedia projector, PABX, photocopiers, video/audio conferencing. recorded less than 40% availability and functionality. Furthermore, more than 40% of the health professionals claimed never to have used multimedia projector (63.2), video camera (61.5%), fax machine (62.6%), photocopier (52.9%), video recorder (51.7%), mobile/land phones (53.4%), video conferencing (70.1%) and Internet searches (41.4%) in the course of their professional duties. While those ICTs used very often include digital camera (53.4%), computers (61.5%), email (40.2%). Health professionals in Ilorin metropolis claimed that some of the ICT facilities they use are available on personal purchase.

Further analysis to correlate the availability and use of ICTs among health professionals in public and private hospitals revealed a significant difference in the use of email, computers, projectors and Internet searches respectively. With a p-value of 0.000, private hospitals recorded higher percentage in the availability and use of computers, projectors, Internet and emails. On the other hand, there was no significant difference in the availability and use of other ICT equipment and enabled services among health professionals in public and private hospitals.

Task and activities carried out with ICT

Despite the fact that the respondents are qualified specialists and well educated and should be able to understand the relevance of ICT services for their research and day to day operations, it is surprising to note that more than half of the health professionals claimed not to use ICTs for research (62.6%), collaboration (88.5%), skills acquisition (75.9%) data handling (71.8%) and training (79.3%). But more than half claim to use ICTs for hospital work (61.5%). This is shown in table 3 below:

Table 3: Activities ICTs are used for

	Yes		No	
	Freq	%	Freq	%
Research	65	37.4	109	62.6
Hospital work	107	61.5	67	38.5
Data handling	47	27.0	127	73.0
Training	36	20.7	138	79.3
Personal Work	76	43.7	98	56.3
Collaboration	20	11.5	154	88.5
Skills acquisition	42	24.1	132	75.9

As shown in table 4, some type of information accessed with ICTs recorded more than 40%. These are clinical information (48.3%) and health information (66.1%), while those that recorded less than 40% are research information (30.5%), administrative information (23.0%), educational information (27.6%) and career information (11.5%).

Table 4: Type of Information Assessed using ICT Enabled Service

	Yes		No	
	Freq	%	Freq	%
Clinical information	84	48.3	90	51.7
Health information	115	66.1	59	33.9
Research information	53	30.5	121	69.5
Administrative information	40	23.0	134	77.0
Educational information	48	27.6	126	72.4
Career information	18	11.5	154	88.5

Access to ICTs

More than half of the health professionals (54.5%) started using ICTs less than five years ago, while others (45.5%) started using ICTs more than five years ago. Also, the health professionals made use of ICTs in their offices, cyber cafés, homes and schools respectively, health professionals' use of ICTs in the office recorded the highest percentage (49.4%).

Table 5: Places ICTs are made use of

	Yes		No	
	Frequency	%	Frequency	%
Office	86	49.4	88	50.6
Cyber café	60	34.5	114	65.5
Home	67	38.5	107	61.5
School	18	10.3	156	89.7

Findings also revealed that very few health professionals attended formal training (31.6%), workshops/seminars (32.8%) and conferences (12.1%) on ICT use. 29.3% claimed they have never attended any training on ICT use, 20.7% attended more than two years ago, while the remaining 50% attended one event or the other on ICT use in the last two years. For the respondents who had never attended any event focusing on the use of ICT facilities, the reasons for never having attended include lack of budget (14.4%), time constraint (14.4%) and lack of courses at proximity (12.6%). When asked to rate their skill level with the use of ICT, more than half of the health professionals rated themselves from good to excellent as shown in table 5.

Table 6: Skill level with the use of ICT

		Frequency	Percent
	Poor	7	4.0
	Fair	69	39.7
	Good	70	40.2
	Very good	20	11.5
	Excellent	8	4.6
	Total	174	100.0

Impact of ICTs

Table 7 below shows the ways health professionals feel ICTs have impacted on their job functions and productivity. Majority (56.3%) of the health professionals claimed that ICTs have impacted positively on their job functions and productivity in areas such as increased efficiency, better doctor – patient rapport, quicker medical diagnosis and easy collaboration among health professionals. In another way, the health professionals in Ilorin metropolis claimed that ICTs have impacted on their job functions and productivity is in the area of seeking medical advice on modern surgical skills like telemedicine.

Table 7: Ways ICTs have Impacted on Health Professionals' Job Functions and Productivity

	Poor		Fair		Good		Very Good		Excellent	
	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%
Increased efficiency	3	1.7	25	14.4	63	36.2	34	19.5	17	9.8
Quicker medical diagnoses	12	6.9	29	16.7	48	27.6	48	27.6	24	13.8
Better doctor – patient rapport	2	1.1	29	16.7	78	44.8	24	13.8	13	7.5
Increase number of publications	22	12.6	28	16.1	51	29.3	23	13.2	11	6.3

More exposure to scrutiny from patients	25	14.4	30	17.2	56	32.2	19	10.9	18	10.3
Has helped to facilitate remote consultation, diagnosis & treatment	14	8.0	30	17.2	51	29.3	42	24.1	17	9.8
Easy collaboration among health professionals	17	9.8	16	9.2	54	31.0	41	23.6	14	8.0
Enabled access to relevant medical training	4	2.3	32	18.4	44	25.3	50	28.7	25	14.4

Attitude towards ICTs

Table 7 shows the frequency distribution of the extent of agreement or disagreement to some statements about ICTs. The Health Professionals were asked to select from a list of options on whether they strongly agree, disagree or strongly disagree with some statements. As revealed in the table, more health professionals agree and strongly agree to the positive statements about ICTs like facilitating access to patient information and providing new tools for sharing knowledge. Likewise, a high percentage of the health professionals disagree and strongly disagree to the statement that they completely lack confidence in ICTs.

Table 8: Knowledge of Relevance of ICT

	SA		A		D		SD	
	q	%	q	%	q	%	q	%
ICTs can cut administration costs	7	3	5	6	5	4	0	7
ICTs can facilitate access to patient information	8	6	9	7	7	0	5	9
ICT applications provide new tools for sharing knowledge	9	1	4	8	9	2	5	9
Network services, including e-mail and multimedia conferencing, are essential lifelines for medical practitioners	0	5	6	2	8	6	5	9
I feel very confident using ICTs	2	6	0	2	2	4	3	7
I feel I can cope with the use of ICTs	2	4	1	1	8	1	5	9
I am completely lacking in confidence	6	4	6	7	7	3	7	0

Majority of the health professionals are of the view that ICTs are relevant to the health profession in many ways. They agree that ICTs will deliver great benefits to medical practitioners (92.0%); every health worker must be equipped with ICTs (86.3%); the role of ICTs in health research is very essential (89.1%); ICTs can be informative for scaling up medical projects (84.5%); ICTs can link with and inform social and

interpersonal communication practices that focus on health (82.8%). On the other hand, a greater part of the health professionals disagree that ICTs will never play an important role in medicine (84.5%) and that ICTs are encroaching on medicine (83.9%).

Also, majority of the health professionals believed that organizing training regularly (94.3%), provision of enough ICT equipment/services (94.3%), implementing an effective health information system (96.0%), easy access to the Internet (94.8%), adequate power supply (93.7%), and maintenance of equipment (95.4%) among others will enhance ICT use among health professionals. Apart from this, they proposed that giving awards and scholarships to the best ICT medical practitioners in Ilorin will enhance ICT use.

Constraints to the Use of ICTs

As revealed in table 9 below, about 50% of the health workers believe that the factors that constitute barriers to the use of ICTs to a high extent are lack of proper training, lack of ICT facilities, lack of electricity supply and lack of physical access. On the other hand, more than 50% of the health professionals identified those factors that constitute barrier to a low extent which are insufficient knowledge on use, lack of affordability and security/privacy issues.

Table 9: Factors that Constitute Barrier to ICT Use

	Low		Medium		High	
	Freq	%	Freq	%	Freq	%
Lack of proper training	57	32.8	37	21.3	80	46.0
Lack of ICT facilities	47	27.0	33	19.0	94	54.0
Insufficient knowledge on use	107	61.5	28	16.1	39	22.4
Lack of time	75	43.1	54	31.0	45	25.8
Lack of electricity supply	42	24.1	50	28.8	82	47.0
Lack of physical access	49	28.2	37	21.3	88	50.6
Lack of affordability	92	52.9	49	28.2	33	18.9
Failure of equipment	83	47.7	41	23.6	50	28.8
Security/privacy issues	100	57.5	43	24.7	31	17.8

5.0 Discussion of Results

This study has clearly demonstrated the place of ICTs on the job functions and productivity of health professionals in public and private hospitals in Ilorin metropolis due to the numerous tasks performed through the use of these tools and more so, a positive attitude towards usage by the health professionals. The results from this study are discussed below in line with the research objectives:

(a) Determining the state of ICT facilities in public and private hospitals.

The first objective of this study was to determine the state of ICT facilities in public and private hospitals. The result of the analysis as revealed in table 1 shows that the available and functioning ICT facilities utilized by the health professionals such as TV Set (87.4%), DSTV (66.1%), Computers (75.9%), Printer (62.6%), Mobile phones (82.8%), among others, were commonly used in these hospitals. On the other hand, non use and non availability of ICT facilities such as video conferencing (70.1%), multimedia projector (63.2%), and video camera (61.5%) recorded a high percentage due to high cost of procuring these facilities by the hospitals. This accounted for the minimal usage of Internet searches (43.1%) and email (50.6%) respectively. On the whole, the available and functioning ICTs in these hospitals were the commonly affordable ones. Comparatively, a low level of usage of Internet driven facilities was reported by the health professionals; thereby hampering the delivery of quality health services and access to current health information. The low availability of Internet driven facilities is peculiar to past studies on ICT usage by health professionals (Idowu et al 2003; Adeyemi and Adebogun, 2004; Adedoyin, Imam and Oladapo 2009).

Furthermore, the significant difference recorded in favor of health professionals in private hospitals as far as the availability and use of computers, Internet, emails and projectors are concerned, is understandable. Factors such as superior service delivery, competitive edge, improved customer service and best practices among others, contribute to such notable differences. The poor maintenance culture of government in Nigeria which has rendered many of the available equipment non functional is also another negative factor. According to Castells (2000), there is concern about examples of waste, delay, mismanagement and corruption within the public sector in Africa, all of these cause inefficiency in the conversion of public expenditure into public services.

(b) Determining task and activities performed with ICT tools

The second objective was to determine the task and activities performed with ICT tools. The result of the analysis as revealed in table 2 showed that a high proportion of the health professionals were less involved in ICT related activities such as research (62.6%), collaboration (88.5%), skills acquisition (75.9%) data handling (71.8%) and training (79.3%). On the other hand, more than half of the respondents claimed to use ICTs for hospital work majorly (61.5%). This depicts that

most of the health professionals were either preoccupied with their daily task of attending to patients or attached to other mundane tasks. Health professionals need to create time to explore the benefits of ICTs through research, skill acquisition, collaboration and so on. This accounted for high percentage of those that needed health information (66.1%) compared to other types of information (research, clinical, career, and administrative) respectively. It is evident here that health professionals in Kwara central senatorial district collaborate only with their colleagues that are within close range and therefore do not participate in transnational exchange of experiences since this can only be done via Internet enabled services which most of the hospitals lack.

(c) Determining the places of access to ICTs

The third objective of this study was to determine the places of access to ICTs by the health professionals. Table 4 revealed that more than half of the health professionals (54.5%) had access to ICTs less than five years ago from various places such as cybercafés, homes, schools and offices. The demographic profile of the health professionals further revealed that (48.9%) had few years of professional experience between 1-5 years which is a determinant of the number of years of access to ICTs. Out of the identified places of access, the office was the first point of access to the use of ICTs. This shows that hospitals should be better equipped with ICTs equipment and infrastructures because that is the first point of access for health professionals.

(d) Impact of ICTs on their job functions and productivity

The fourth objective determined the impact of ICTs on health professionals' job functions and productivity. Table 6 revealed that ICTs have positively impacted on the job functions and productivity of health professionals in areas such as increased efficiency, better doctor – patient rapport, quicker medical diagnosis and easy collaboration among others. This finding supports Davies et. al, (2006) in the identification of personal, interpersonal and organizational benefits of ICTs by health professionals.

(e) Evaluating the attitude of health professionals to the use of ICTs

The fifth objective evaluated the attitude of health professionals to the use of ICTs. In table 7, it was revealed that majority of the health professionals exhibited a positive attitude towards the use of ICTs due to the high level of awareness and usefulness of ICTs to the medical profession. Majority of the health professionals affirmed positively to the following statements: Health professionals should be equipped with ICTs (86.3%), ICTs in health research is very essential (89.1%); ICTs can be informative for scaling

up medical projects (84.5%); ICTs can link with and inform social and interpersonal communication practices that focus on health (82.8%). On the other hand, majority of the health professionals disagree that ICTs will never play an important role in medicine (84.5%) and that ICTs are encroaching on medicine (83.9%). This result has clearly shown that ICTs are indispensable tools to the modern day health professionals because of its numerous benefits that include the use and the delivery of quality health care services. This is in line with the findings of Adedoyin, Imam and Oladapo (2009) on ICT literacy of health professionals which revealed that majority of health professionals in developing countries are becoming aware of the usefulness of ICTs to the medical profession.

(f) Constraints to the utilization of ICTs

The Sixth objective determined the constraints to the utilization of ICTs by the health professionals. As revealed in table 8, barely half (50%) of the health professionals identified barriers of high extent such as lack of proper training, lack of ICT facilities, lack of electricity supply and lack of physical access. On the other hand, barriers of low extent constituted insufficient knowledge on use, lack of affordability and security/privacy issues. This result confirms the findings of past studies on the common barriers militating against the use of ICTs by health professionals in developing countries (Idowu et. al, 2003; Ouma and Herseman, 2008; Chandrasekhar and Ghosh, 2001). However, these constraints to the utilization of ICTs can be minimized through the development and proper enforcement of ICT usage that must be enforced on health professionals in public and private hospitals by the relevant regulatory agencies.

6.0 Conclusion and Recommendations

This study has clearly revealed that although health professionals in the study area are aware of the usefulness of ICTs to the medical profession, however, access, usage and volume of ICTs are low. The health professionals are handicapped because there are some core ICT enabled services like video/audio conferencing that are not available for use within the hospital environment and which they cannot own by themselves. Even though the health professionals can afford to provide for themselves other ICT enabled services like Internet searches, provision of ICT equipment and services by hospitals is very fundamental to achieving the desired impact of ICTs in the health sector. The lack of necessary ICT facilities in these hospitals have resulted in low access to adequate and reliable information. This is hampering efficient health care delivery services.

Having identified barriers to the superior use of ICTs for enhanced access to current and relevant medical information, the following recommendations are suggested:

1) Hospital regulatory bodies should collaborate with International agencies in the area of funding by giving out grants for the successful implementation of sophisticated ICTs in health institutions in public and private sectors.
2) There should be training and re-training of health professionals on ICT use so that health professionals can be abreast of the latest happenings in the ICT world. Also, it is not enough to make available state of the art ICT equipment; there must be training on the use of such ICTs to guarantee adequate technical know-how and proper maintenance.
3) Internet access within hospital environments should be given high priority. With Internet access, health workers can collaborate with their colleagues outside the Ilorin metropolis and share experiences that will impact positively on their duties.
4) Hospitals should have an in-house policy that mandates staff to use ICTs in the discharge of their duties. This is because there are bound to be health professionals, particularly the older ones, that will be carefree when it comes to adopting some specific technologies.
5) Proper maintenance of ICT equipment should be done at all times to ensure that all available ICTs are functioning and serving.
6) The Government, in collaboration with hospitals' regulatory bodies, should put in place policies that will ensure that all health institutions have a minimum standard of ICT facilities within their hospital premises.
7) Electricity is an infrastructural necessity in the use of ICTs, hence Government should give it the priority it deserves. Also, all hospital management should supplement government electricity supply in their various hospital environments with modern power generation technologies like the solar and gas energy plants.

References

Adeyemi, A., & Ayegboyin, M. (2004, March). A study on the use of information systems to prevent HIV/AIDS in Lagos state, Nigeria. Paper presented at Informedica: Information and Communication Technologies in Healthcare Development. 3rd Virtual Congress in Internet.

Adeyoyin, S., Imam, A. and Oladapo, Y. (2009). Health workers' ICT literacy in a Nigerian University Teaching

Hospital. *The Information Technologists,*Vol. 6 (2).

Castells, M. (2000). End of Millennium, 2nd edition. Oxford: Blackwell.

Chandrasekhar, C. P., & Ghosh, J. (2001). Information and Communication Technologies and Health in low-Income Countries: The Potential and the Constraints. *Bulletin of the World Health Organization*, *79*, 850–855.

Davies, J., Trude, B., Mcconnell, H., Ramirez, R.,Shields, T., Drury, P., Kumekawa, J., Louw, J., Fereday, G., and Nyamai-Kisia, C. (2006). Improving Health, Connecting people: The Role of ICTs in the Health Sector of Developing countries. A framework paper commissioned by InfoDev program.

Electronic Government (2004). South African Health Example 2, (1), 31. Available at: www.hisp.org. Accessed on 10/02/2013.

Empirica (2008) Benchmarking ICT use among General Practitioners in Europe. Available at: http://www.rcc.gov.pt/SiteCollectionDocuments/ICT_Europe_final_report08.pdf. Accessed on 13/03/2013.

Gatero, G. (2011). Utilization of ICTs for Accessing Health Information by Medical Professionals in Kenya: A Case Study of Kenyatta National Hospital. *Journal of Health Informatics in Developing Countries,*Vol. 5 (1), *60-88.*

Godlee F. et al (2004) Can we achieve health education for all by 2015? Lancet, 364: 295 – 300. Available at: *http://imagethelancet.com/extra/04art6112web.pdf.* Accessed on 13/03/2013.

Greenberg, (2005). ICTs for Poverty Alleviation: Basic tool and Enabling Sector. Stockholm: SIDA.

Idowu, B., Ogunbodede, E. and Idowu, B. (2003). Information and Communication Technology in Nigeria : The Health Sector Experience *Journal of Information Technology Impact, 3 (2), 69-76.*

Jimoh L., Pate, M. A., Lin, L. and Schulman, K. A. (2012). A model for the adoption of ICT by health workers in Africa. *International Journal of Medical Information*, 81, (11), 773-81.

Olatokun, W. M. and Adeboyejo, O. C. (2009). Information and Communication Technology Use by Reproductive Health Workers in Nigeria: State of the Art, Issues, and Challenges. *Human Technologists*, 5 (2), 181–207.

Ouma, S. and Herselman, M.E. (2008). E-health in Rural Areas: Case of Developing

Countries. *World Academy of Science Engineering and Technology*, 40, 560- 566.

United Nations Development Programme (2007). E-Health Tools for the Asia Pacific Region. Available at: www.apdip.net/apdipenote/22.pdf. Accessed on 11/04/2013.

APPENDIX

QUESTIONNAIRE

Dear Respondent,
We are conducting a research on "Availability, Accessibility and Frequency of Use of ICT Tools by Health Professionals in Ilorin Metropolis". I am glad to inform you that you have been chosen to participate in this study with regards to the above subject.

Your responses and views are highly important in achieving the objectives of this study. The information provided will be treated as private and confidential and will be used for the purpose of this study only.
Thank you for your valuable time, attention and cooperation.

Yours Faithfully,
Researchers

SECTION A - DEMOGRAPHIC INFORMATION.

1. Gender: Male ( ) Female ( )

2. Age: 26 – 30yrs ( ) 31 – 35yrs ( ) 36 – 40yrs ( ) 41 – 45yrs ( )
46 – 50yrs ( ) 51 and above ( )

3. Type of Health Institution: Public ( ) Private ( )

4. What category of health worker are you?

Medical Doctor ( ) Nurse ( ) others (please specify): _____

5. Highest Educational Qualification

Registered Nurse ( )

Registered Nurse Midwife ( )

B. Sc Nursing ( )

M. Sc Nursing ( )

PhD Nursing ( )

MBBS ( )

Additional degrees in other areas ( )

Others (please specify) _____

6. Years of professional experience:
0-5 ( ) 6-10 ( ) 11–15 ( ) 16-20 ( ) 21-25 ( ) 26–30 ( )
31 and above ( )

SECTION B – ICT Facility Used and Extent of Use

7. Which of the following ICT equipments are available in your Hospital?

ICT products and equipment	Available & Functioning	Available but Not Functioning	Not Available	Don't Know
Multimedia Projector				
PABX (Intercom)				
Video Camera				
Fax machine				
DSTV				
Photocopier				
Video Recorder				
T.V sets				
Digital camera				
Mobile phone/Nitel phone				
Computers (Desktop/Laptop)				
Printers				
Scanners				
Others (specify)				
ICT enabled services				
email				
WWW				

Videoconferencing				
Audio conferencing				
Internet searches e.g. Google				
Others (specify)				

8. Which of the following ICT facilities do you make use of in your work and how often?

ICT products and equipment	Never	Rarely	Often	Very often
Multimedia Projector				
PABX (Intercom)				
Video Camera				
Fax machine				
DSTV				
Photocopier				
Video Recorder				
Scanners				
T.V sets				
Digital camera				
Mobile phone/Nitel phone				
Computers (Desktop/Laptop)				
Printers				
Others (specify)				
ICT enabled services				
email				
WWW				
Videoconferencing				
Audio conferencing				

Internet searches e.g. Google				
Others (specify)				

SECTION C - Task and activities carried out with ICT

9. Which of the following activities do you use ICT for?
Research ( ) Hospital work ( ) Data handling ( ) Training ( )
Personal Work ( ) Collaboration ( ) Skills acquisition ( )
Others (please specify) _____

10. What type of health related information do you access using ICT enabled services?
Clinical information ( ) Health information ( )
Research information ( ) Administrative information ( )
Educational information ( ) Career information ( )
Others (please specify) _____

SECTION D - Access to ICTs

11. How long ago did you **first** start to use ICT? (Please answer in *either* months or years) _____ Months ago _____ Years ago

12. Which of the following places do you usually make use of ICT?
Office ( ) Cyber cafés ( ) Home ( ) School ( )
Others ( )

13. Have you ever attended any of the following events on ICT use?
Formal Training ( ) Workshops/Seminars ( ) Conferences ( )
Others (please specify) _____

14. When last did you attend any of the above, focusing on the use of ICT facilities?
1-6 months ago ( ) 6-12 months ago ( ) 1-2 years ago ( )
More than two years ago ( ) Never ( )

15. If your answer to 14 above is Never, which of the following has been the reason?
Lack of budget for staff training courses ( )
Pressure of current workload ( )
No courses at any location nearby ( )
Personal family commitments ( )
Lack of interest ( )
Time constraint ( )
Others (please specify) ----------

16. How would you rate your skill level with the use of ICT?
Poor ( ) Fair ( ) Good ( ) Very Good ( ) Excellent ( )

SECTION E - Impact of ICTs

17. In what ways have ICTs impacted on your job functions and productivity?

Increased efficiency	Poor	Fair	Good	Very Good	Excel lent
Quicker medical diagnoses					
Better doctor – patient rapport					
Increase number of publications					
More exposure to scrutiny from patients					
Has helped to facilitate remote consultation, diagnosis & treatment					
Easy collaboration among physicians					
Enabled access to relevant medical training					
Others (specify):_____ _____					

SECTION F - Attitude towards ICTs

18. Please indicate the degree to which you agree or disagree with the following statement by circling SA for strongly agree, A for agree, D for disagree, or SD for strongly disagree.

	SA	A	D	SD
ICTs can cut administration costs				
ICTs can facilitate access to patient information				
ICT applications provide new tools for sharing knowledge				
Network services, including e-mail and multimedia conferencing, are essential lifelines for medical practitioners				
I feel very confident using ICTs				
I feel I can cope with the use of ICTs				
I am completely lacking in confidence				

19. What are your views on the relevance of ICTs to the Health profession?

I expect that ICTs will deliver great benefits to medical practitioners ( )

I believe every health worker must be equipped with ICTs ( )

The role of ICTs in health research is very essential ( )
ICTs can be informative for scaling up medical projects ( )
ICTs can link with and inform social and interpersonal communication practices that focus on health ( )
I don't think that ICTs will ever play an important role in medicine ( )
I am worried about the way in which ICTs are encroaching on medicine ( )

SECTION G - Constraints to the Use of ICTs

20. Are ICT facilities readily available to you in your hospital?
Very Available ( ) Available ( ) Fairly Available ( )
Not Available ( )

21. Do any of the following reasons constitute a barrier to your using ICTs?

	Not at all	Low	Medium	High
Lack of proper training				
Lack of ICT facilities				
Insufficient knowledge on use				
Lack of time				
Lack of electricity supply				
Lack of physical access				
Lack of affordability				
Failure of equipment				
Security/privacy issues				
Others (please specify)				

22. What do you think can be done to enhance ICT use among Medical Practitioners?

Organize training regularly	Yes	No
Provision of enough ICT equipment/services		
Implementing an effective health information system		
Easy access to the internet		
Adequate power supply		
Maintenance of equipments		

Free access to online medical journals		
Provision of latest ICT equipments		
Access to electronic publishing		
Having and implementing ICT strategy and policy		
Others (please specify)		

SOME ASPECTS OF TOPOLOGICAL SORTING

[1]Yohanna Tella (M.Sc, Ph.D),

[2]D. Singh (M.Sc, Ph.D)
&

[3]J.N. Singh (M.Sc, Ph.D)

[1] (Corresponding Author)
Department of Mathematical Sciences
Kaduna State University, Kaduna, Nigeria.
Email: yot2002ng @kasu.edu.ng

[2] (Former Professor, Indian Institute of Technology Bombay).
Department of Mathematics
Ahmadu Bello University, Zaria, Nigeria.
Email: mathdss @yahoo.com

[3]Department of Mathematics and Computer Science, Barry University,
11300 North East Second Avenue
Miami Shores, FL 33161, USA

Abstract: In this paper, we provide an outline of most of the known techniques and principal results pertaining to computing and counting topological sorts, realizers and dimension of a finite partially ordered set, and identify some new directions.

***Key words*:** Partial Ordering, topological sorting, realizer, dimension.

1. Introduction

Topological sorting has been found particularly useful in sorting and scheduling problems such as *PERT charts* used to determine an ordering of tasks, *graphics* to render objects from back to front to obscure hidden surfaces, *painting* when applying paints on a surface with various parts and identifying errors in *DNA* fragment assembly (Knuth, 1973; Skiena, 1997; Rosen, 1999). In this paper, we provide systematically an outline of most of the techniques and principal results pertaining to computing and counting topological sorts, realizers and dimension of a partially ordered set (Poset), and identify some new directions.

2. Definitions of some terms and related results pertaining to partially ordered sets

We borrow these definitions from various sources (Brualdi *et al*., 1992; Jung, 1992; Trotter 1991). Let $(P, \leq_p)$ denote a finite partially ordered set along with an implicit assumption that P denotes the underlying set and $\leq_p$ denotes its order relation. Moreover, $\leq$ stands for reflexive partial order and $<$ for strict partial order.

For an element $a \in P$, the set $U(a) = \{x \in P | a <_p x\}$ is called an *open upset* of a .The set $U[a] = \{x \in P | a \leq_p x\}$ is called a *closed upset* of a. For any nonempty subset $A \subseteq P$, $U(A) = \{x \in P | a <_p x \;\; a \in A\} = \bigcup_{a \in A} U(a)$ and $U[A] = \{x \in P | a \leq x, \; a \in A\} = \bigcup_{a \in A} U[a]$.

Similarly, the open and closed down sets can be defined. Note that the closed upset and the closed down set are also called *filter* and *ideal* respectively.

An element $y \in P$ is said to cover an element $x \in P$ if $x \neq y$ and $x \leq_p y$ with no element $z \in P$ such that $x \leq_p z \leq_p y$. Sometimes we also say that y *is an immediate successor of* x or x *is an immediate predecessor of* y. For every$x, y \in P$, if $x \leq_p y$ then the pair (x, y) is said to be a *comparability* of P. Two elements $a, b \in P$ are called comparable, denoted $a \perp b$ or $a|b$ if either $a \leq_p b$ or $b \leq_p a$; and incomparable denoted $a||b$, if both $a \not\leq_p b$ and $b \not\leq_p a$. Also, $a <_p b$, iff $a \leq_p b$ and $a \neq b$.

The incidence of a poset $(P, \leq_p)$, denoted $inc(P)$, is defined as the set $\{(x, y) \in P \times P : x||y \text{ in } \leq_p\}$.

A pair $(x, y) \in inc(P)$ is called a *critical pair* if

$u \leq_p x$ in $\leq_p$ implies $u \leq_p y$ in $\leq_p$ and

$v \;_p\!\geq y$ in $\leq_p$ implies $v \;_p\!\geq x$ in $\leq_p$ for all $u, v \in P$.

Also, the set of all critical pairs is denoted $crit(P)$.

A subset A of a poset $(P, \leq_p)$ is called a *subposet* if the *suborder* $\leq_A$ defined on A, is the restriction of $\leq_p$ on pairs of elements of A. In other words, a subposet is a subset of $(P, \leq_p)$ with the induced order.

A chain in P is a subposet of P which is a linear order. The length of a chain C of P is $|C| - 1$. An antichain in P is a subset of P containing elements that are mutually incomparable.

Two posets $(P, \leq_p)$ and $(Q, \leq_Q)$ are *Isomorphic*, $P \cong Q$, if there exists order preserving bijection $\varphi : P \to Q$ such that $x \leq_p y \Leftrightarrow \varphi(x) \leq_Q \varphi(y)$.

A poset of the type $(\{a, b, c, d\}, < | a < c, \;\; b < c, b < d)$ is called an N-poset as its Hasse diagram looks like the letter N:

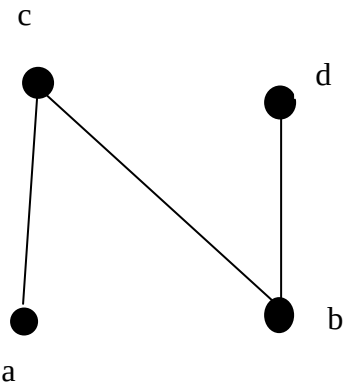


Figure 1

A poset $(P, \leq_P)$ is called *N-free* if there exists no subposet A of P isomorphic to an N-poset.

Let $(P, \leq_P)$ and $(Q, \leq_Q)$ be two disjoint posets. The *disjoint* (*cardinal*) *sum* $P + Q$ is the poset $(P \cup Q, \leq_{P+Q})$ such that $x \leq_{P+Q} y$ if and only if $x, y \in P$ and $x \leq_P y$ or $x, y \in Q$ and $x \leq_Q y$. The *linear* (*ordinal*) *sum* $P \oplus Q$ is the poset $(P + Q, \leq_{P \oplus Q})$ such that $x \leq_{P \oplus Q} y$ if and only if $x \leq_{P+Q} y$ or $x \in P$ and $y \in Q$ with x preceding y. In other words, $P \oplus Q$ is obtained from $P + Q$ by adding $x \leq y$ (or x preceding y) for all $x \in P$ and $y \in Q$.

A k- antichain is defined as the disjoint union of k singletons.

A poset $(P, \leq_P)$ is called *series-parallel* if it can be constructed from singletons by using disjoint union and linear sum.

A crown on $2n$ elements $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ is the partial order defined:

For all indices $i \neq j$, the elements a_i and a_j are incomparable, the elements b_i and b_j are incomparable, but $a_i < b_j$; and for each i, the elements a_i and b_i are incomparable.

A 4-crown (or a crown with four elements) poset is isomorphic to **2** $\oplus$ **2** (where **2** is a two elements antichain) and is a series parallel poset**.** The *N*-poset can be described as

2 $\oplus$ **2** with one comparability missing. For example let $P = \{x, y\}$ and $Q = \{r, s\}$ be the two elements antichains. It is clear that $x <_{P \oplus Q} r, x <_{P \oplus Q} s, y <_{P \oplus Q} r, y <_{P \oplus Q} s$ for the poset $P \oplus Q$ (by definition).

The Hasse diagram follows:

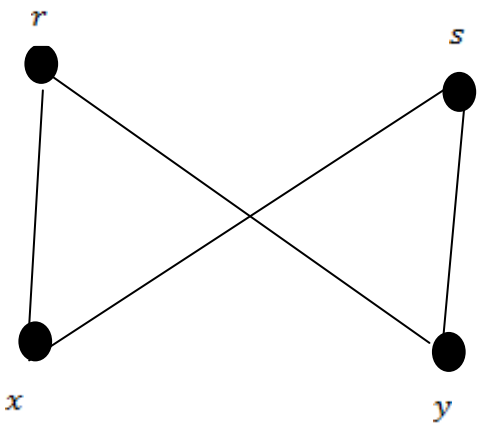
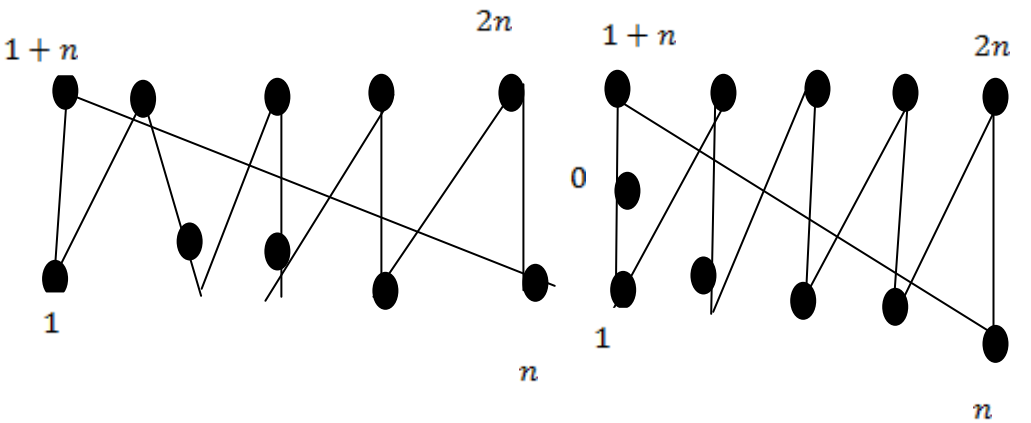


Figure 2

figure 2 above is an N poset Hasse diagram with the comparability $x <_{P \oplus Q} s$ missing.

Series-parallel posets can also be characterized as N-free posets (Valdes *et al.,1982*)

A *cycle* is a poset $(P, \leq_P)$ with Hasse diagram in figure 3(a) where $n \geq 2$



(a) **(b)**

Figure 3 (a) Cycle (b) Crown

The *young's lattice* $L(m,n)$, where m,n are positive integers, is a poset defined on $\{(a_1, a_2, \ldots, a_m) | 0 \leq a_1 \leq a_2 \leq \cdots \leq a_m \leq n, a_i \in \mathbb{Z}\}$ with the order relation:

$(a_1, a_2, \dots, a_m) \leq (b_1, b_2, \dots, b_m)$ if and only if $a_i \leq b_i, i = 1,2, \dots, m$.

The *height* of a poset $(P, \leq_p)$, denoted $h(P)$, is defined to be the cardinality of its longest chain. The width of $(P, \leq_p)$, denoted by $w(P)$, is defined to be the cardinality of its largest antichain. It is easy, though not trivial, to see that the following results hold

(Dilworth, 1950; Brualdi *et al.*, 1992):

$w(P)$ equals the minimum number of chains in a partition of P into chains, $h(P)$ equals the minimum number of antichains in a partition of P into antichains,

$w(P) \geq w(P \setminus \{x\}) \geq w(P) - 1$

and

$h(P) \geq h(P \setminus \{x\}) \geq h(P) - 1$.

A Poset P is called *width-critical* if $w(P \setminus \{x\}) < w(P)$ and *height-critical* if $h(P \setminus \{x\}) < h(P)$, for all $x \in P$.

It follows that a poset P is width-critical if and only if P is an antichain and height-critical if and only if P is totally ordered.

3. Algorithms for constructing linear extensions

3.1 Definitions of some basic terms and related results pertaining to ordered sets

Let $(P, \leq_p)$ be a finite nonempty poset. A total ordering $<$ is said to be compatible with the partial ordering $\leq_p$ if $a < b$ whenever $a \leq_p b$. The scheme for constructing a compatible total ordering from a partial ordering is called *topological sorting* and the outcome is called a *topological sort (*or *linear extension).* In other words, a linear extension of P is a linear order which contains P.

Let L denote a linear extension of $(P, \leq_p)$ and $L(P)$ denotes set of all linear extensions of P. $L(P)$ is nonempty for any P (Szpilrajn, 1930). That is, every order can be extended to a linear order. In fact, a stronger result has been proved: Let $\leq_p$ be an order on P and let $x, y \in P$ such that $x||y$. Then there exist two linear extensions $\leq_{L_1}$ and $\leq_{L_2}$ of $\leq_p$ such that $x \leq_{L_1} y$ and $y \leq_{L_2} x$.

A linear extension L is said to *reverse* the incomparable pair (x, y) when $x \geq y$ in L. A family $\mathcal{R}$ of linear extensions of P reverses $A \subseteq inc(P)$ if for every $(x, y) \in A$, there exists some $L \in \mathcal{R}$ such that $x \geq y$ in L.

The dual of a linear extension L of a poset $(P, \leq_p)$ denoted L^d, is a linear order obtained by reversing the order of the linear extension L. The dual of a poset $(P, \leq_p)$, denoted P^d, is the poset obtained by reversing its order.

A consecutive pair (x_i, x_{i+1}) of elements in L is called a *jump* or *setup* of $(P, \leq_p)$ in L if x_i and x_{i+1} are incomparable in $(P, \leq_p)$. We denote the number of jumps of $(P, \leq_p)$ in L by $r(L, P)$. The *jump number* $r(P)$ of $(P, \leq_p)$ is the minimum of $r(L, P)$ over all linear extensions L of $(P, \leq_p)$.

A Poset P is called *jump critical* if $r(P \setminus \{x\}) < r(P)$ for each $x \in P$.

A jump-critical Poset $(P, \leq_p)$ with jump number m has atmost $(m+1)!$ elements (El-Zahar & Schmerl, 1984), and there are precisely 17 jump-critical posets with jump number atmost 3 (El-Zahar & Rival, 1985). It is recognized that characterizing jump-critical posets turns out to be a considerably complicated problem. Pulleyblank proved that jump number problem viz. *schedule the tasks to minimize the number of jumps is* NP*-hard* (Bouchitte & Habib, 1987).

It follows from Dilworth's theorem that $r(P) \geq w(P)$. If $r(P) = w(P)$, then $(P, \leq_p)$ is called a *Dilworth poset* or simply a D- poset . It is shown that a poset which does not have a subposet isomorphic to a cycle in figure 2(a) is a D- poset (Duffus et al., 1982).

Syslo ((Bouchitte & Habib, 1987). put forward a polynomial algorithm to characterize Dilworth posets in the case where the antichain of maximal elements is a maximal-sized antichain. It is observed that the class of Dilworth posets does not seem to be nice with respect to computational complexity.

If $r(L,P) = r(P)$, then L is called an *optimal linear extension* of $(P, \leq_p)$. We denote the set of all optimal linear extensions of $(P, \leq_p)$ by $\mathcal{O}(P)$.

3.2 Knuth's (Bucket) sorting algorithm

Essentially, the topological sort of a finite partial order P is a linear order $a_1, a_2, \ldots, a_n$ of elements of P such that $i < j$ whenever $a_i < a_j$ in P i.e; x precedes y in the partial ordering implies x precedes y in the linear extension (Knuth, 1973). The idea is to pick a minimal element and then to remove it from the poset, and continue the process with the truncated poset until it gets exhausted. A very fast algorithm and its implementation for computing a topological sort of a poset is presented in (Knuth, 1973) . As a matter of fact, this is a well-documented work on sorting. In its simplest form, the algorithm for constructing a total ordering in the finite poset $(P, <_p)$ can be depicted as below:

Since $(P, <_p)$ is *finite* and nonempty, it has minimal elements. Let a_1 be a minimal element i.e. a_1 is not preceded by any other object in the ordering $(P, <_p)$ which is chosen first. Again $(P - \{a_1\}, <_p)$ is also a poset. If it is non-empty, let a_2 be one of its minimal elements, which is chosen next and continue the process until no element remains to be further chosen. Since P is finite, this process must terminate and give rise to a sequence of elements $a_1, a_2, \ldots, a_n$ along with the desired total ordering defined by $a_1 < a_2 < \cdots < a_n$. The compatibility of the total ordering $<$ with the original partial ordering $<_p$ can easily be verified i.e. by the definition given above, $a <_p b \Longrightarrow a < b$ for all a, b in the ordering. It needs to be constantly

observed that $a < b$ only if a is chosen before b.

Alternatively, for a subset S of P, we denote the set of minimal elements of P restricted to S by $MIN(S)$. The algorithm LIN for computing a linear extension $[a_1, a_2, \dots, a_n]$ of the poset $(P, <_p)$ (Kierstead *et al., 1987*) is defined:

ALGORITHM LIN:

SET R = P, M = MIN(R)

FOR i = 0, ..., n − 1

CHOOSE $a_{i+1} \in M$

SET R = R \ $\{a_{i+1}\}$, M = MIN(R)

END

For any sequence of choices of the points a_{i+1}, the algorithm LIN produces a linear extension of $(P, <_p)$; and every linear extension of $(P, <_p)$ is obtained from LIN by a suitable sequence of choices of a_{i+1}.

Example 1:
Let $P = \{1,2,3,4,8,12\}$ and the partial ordering relation be "divides" denoted by $|$. The scheme to find a compatible total ordering for the poset $(\{1,2,3,4,8,12\}, |)$ may be outlined as follows:
At first stage, 1 is the only minimal element and hence gets selected. Next, we need to select a minimal element of $(\{2,3,4,8,12\}, |)$. At this stage, 2 and 3 are the two minimal elements, we select 3. Next, we need to select a minimal element of $(\{2,4,8,12\}, |)$.

At this stage, 2 is the only minimal element. Next we need to select any minimal element of $(\{4,8,12\}, |)$. Here, 4 is the only minimal element. Next, as both 8 and 12 are minimal elements of $(\{8,12\}, |)$, we select 12. Finally, 8 is left as the last element. The outcome is the total ordering
$1 < 3 < 2 < 4 < 12 < 8$. A linear representation of the above can be depicted as follows: $1 \longrightarrow 3 \sim 2 \longrightarrow 4 \longrightarrow 12 \sim 8$. Another compatible total ordering for the same partial ordering may be constructed as follows: $1 < 3 < 2 < 4 < 8 < 12$. Hence it follows that a compatible total ordering for a given partial ordering may not be unique. In fact, the size of the family of linear extensions of a poset P varies from 1 (if P is a chain) to $n!$ (if P is an n-element antichain). Note, however, that it may not be even possible to construct such a total ordering if loops were present.

3.3 Depth first traversal algorithm

Another algorithm used for the computation of a topological sort is totally based on *depth first traversal* (*Papamanthou, 2004*):

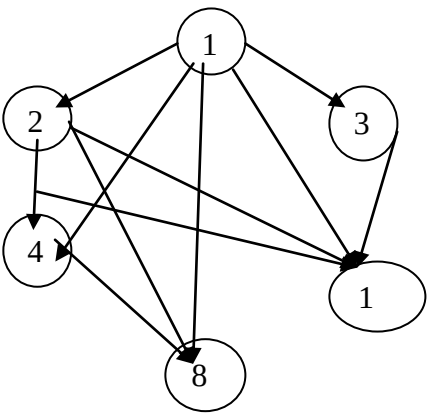
(i) Perform DFS to compute the discovery/finishing times $d[v]/f[v]$ for each vertex v representing an element in the Hasse diagram of the poset.

(ii) As each vertex is finished, insert it to the front of a linked list

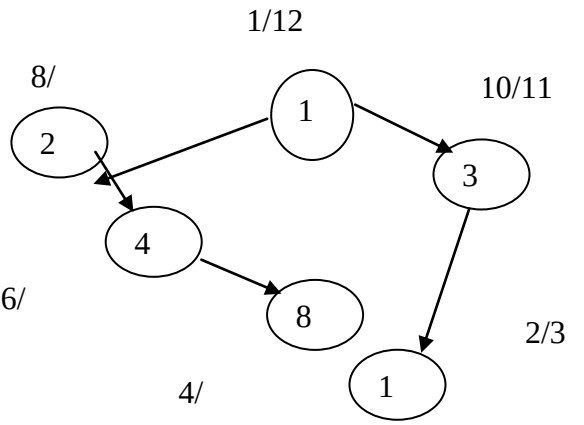
(iii) Return the linked list of vertices

(iv) Output the vertices in reverse order of finishing time to get the topological sort of the poset (Skiena, 1997).

Example 2: Below is an original graph of the poset $P = (\{1,2,3,4,8,12\}, |\ )$ and its Depth first search (DFS) forest.



Original graph of poset P



Depth First Search (DFS) forest of poset P

Figure 4

Final order: $1 < 3 < 2 < 4 < 8 < 12$.

Note that the DFS could generate other distinct topological sorts using the same vertex 1.

4. Algorithm for constructing greedy linear extensions

A more restrictive class of linear extensions of a poset $(P, <_p)$ is obtained by further restricting the choice of a_{i+1} to generate topological sorts called *greedy* linear extensions. The algorithm for computing greedy linear extensions of a poset (*Cogis & Habib, 1979; Brualdi et al.,* 1992) is given as follows:

(i) Choose a minimal element a_1 of P

(ii) Suppose $a_1, a_2, \ldots, a_i$ have been chosen.

T_1: If there is at least one minimal element of

$P \setminus \{a_1, a_2, \dots, a_i\}$ which covers a_i then choose a_{i+1} to be any such minimal

element; otherwise, choose a_{i+1} to be any minimal element of $P \setminus \{a_1, a_2, \dots, a_i\}$.

More precisely, a linear extension of $(P, <_p)$ is greedy if and only if it is obtained from the following algorithm *GREEDY* by a suitable sequence of choices of the points a_{i+1} (*Kierstead et al.*, 1987)*:*

ALGORITHM GREEDY

$SET\ R = P, M = MIN(R), G = M$

$FOR\ i = 0, \dots, n-1$

$CHOOSE\ a_{i+1} \in G$

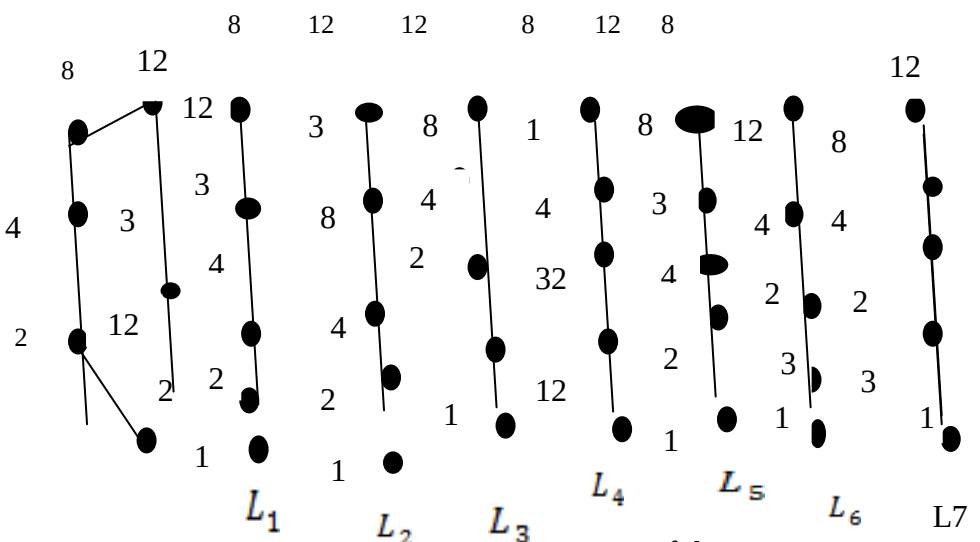
$SET\ R = R \setminus \{a_{i+1}\},\ M = MIN(R)$

$IF\ M \cap U(a_{i+1}) \neq \emptyset$
$THEN\ SET\ G = M \cap U(a_{i+1})$

$ELSE\ SET\ G = M$

END

Example 3: Hasse diagram of the poset $(\{1,2,3,4,8,12\}, |\ )$ and its corresponding linear extensions:



Poset P Linear extensions of the poset $(P, <_p)$

Figure 5

By definition, L_2, L_6 and L_7 are greedy linear extensions of the poset P, but L_1 , L_3 , L_4 and L_5 are not greedy.

Let $\mathcal{G}(P)$ denote the set of all greedy linear extensions of the poset P. A poset $(P, \leq_p)$ is *greedy* if $\mathcal{G}(P) \subseteq \mathcal{O}(P)$; that is, every greedy linear extension is optimal.

Every greedy linear extension is optimal for the jump number on the class of *series parallel* posets (*Cogis & Habib, 1979*) . Every N-free poset is greedy (Rival, 1986). An optimal linear extension of Dilworth poset is necessarily greedy (Bouchitte &Habib, 1987).

The *Young's Lattice* $L(m,n)$ is greedy if and only if one of (1) $m = 1$ or $n = 1$ and (2) $m \leq 2$ and $n \leq 2$ is satisfied. Every poset $(P, \leq_p)$, containing no subposet isomorphic to figure 2(b) given in section 2, satisfies $\mathcal{O}(P) \subseteq \mathcal{G}(P)$ (El-Zahar & Rival, 1985).

A poset P is reversible if $L^d \in \mathcal{G}(P^d)$ whenever $L \in \mathcal{G}(P)$. A poset $(P, \leq_p)$ is reversible if and only if $\mathcal{O}(P) = \mathcal{G}(P)$ (Rival & Zaguia, 1986; Jung, 1992).

5. Algorithm for constructing super greedy (depth-first greedy (dfgreedy)) linear extensions.

A further restrictive class of linear extensions of a poset $(P, \leq_p)$ is the class of *super greedy* (*depth-first greedy (dfgreedy)*) linear extensions.

A greedy linear extension of $(P, \leq_p)$ is *super greedy* if it is obtained by applying the following scheme (*Bouchitte et al.*,1985; *Ducournau & Habib, 1987*) :

(i) Choose for a_1 any minimal element of P

(ii) If $\{a_1, a_2, \dots, a_i\}$ is super greedy, then choose for a_{i+1} any minimal element of

$P \setminus \{a_1, a_2, \dots, a_i\}$ covering a_k, where $k \leq i$ is the greatest subscript, if there exists one; otherwise choose any minimal element of $P \setminus \{a_1, a_2, \dots, a_i\}$.

Alternatively, a linear extension L of $(P, \leq_p)$ is super greedy if and only if it is obtained by applying the following algorithm $SUPER\ GREEDY$ by a suitable sequence of choices of the points a_{i+1}(Kierstead et al., 1987):

$ALGORITHM\ SUPER\ GREEDY$

$SET\ R = P,\ M = MIN(R),\ SG = M$

$FOR\ i = 0, \dots, n-1$

$CHOOSE\ a_{i+1} \in SG$

$SET\ R = R \setminus \{a_{i+1}\}, M = MIN(R), k = i$

WHILE $M \cap U(a\_k) = \emptyset$ AND $k \neq 0$ DO
SET k=k-1

$IF\ k \neq 0,$
$THEN\ SET\ SG = M \cap U(a_k)$

END

In example 3, L_5, L_6 and L_7 are super greedy linear extensions.

The notion of super greedy linear extension was introduced in (Pretzel, 1985), and studied some of its algorithmic properties studied

(Bouchitte et al., 1985) . Every super greedy linear extension is greedy i.e; $\mathcal{SG}(P) \subseteq \mathcal{G}(P)$, where $\mathcal{SG}(P)$ denotes the set of all super greedy linear extensions of a poset $(P, \leq_p)$ (Bouchitte & Habib, 1987). Computational complexity aspect of greedy and super greedy linear extension construction associated with the jump number has been studied (Kierstead, 1986).

6. Counting topological sorts, Dimension, and Realizers of a poset

The following are some established facts in this regard (Trotter, 1991; Brualdi et al., 1992; Skiena , 1997; Schroder, 2003; Kloch, 2007):

6.1 Counting topological sorts

(i) Posets with no elements have exactly one linear extension, the null set.

(ii) A Poset that is a chain has just one linear extension which is itself.

(iii) A Poset that is an antichain of n elements has $n!$ linear extensions

(iv) The number of linear extensions of all other Posets with n elements lies between two bounds mentioned in (ii) and (iii).

(v) $L(P) = \sum_{k\ minimal} L(P \setminus \{k\})$, where $L(P)$ denotes the number of all linear extensions of a given Poset P. The $L(P)$ for various linear extensions outlined above can be computed using the formulae viz;

$$L(P_{greedy}) = \sum_{k\ greedy\ minimal} L(P \setminus \{k\})$$

and

$$L(P_{supgreedy}) = \sum_{k\ super\ greedy\ minimal} L(P$$

where $L(P), L(P_{greedy})\ and\ L(P_{supgreedy})$ denote the number of linear extensions, greedy linear extensions and super greedy linear extensions of a given Poset $(P, <_p)$ respectively. The expressions $L(P), L(P_{greedy}), L(P_{supgreedy})$ are also useful for enumerating all the linear extensions of each kind.

Example 4. We enumerate $L(P), L(P_{greedy})\ and\ L(P_{supgreedy})$ of the poset $(P, <_p)$ given by the Hasse diagram below:

:

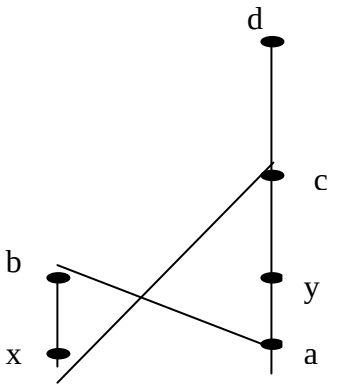


Figure 6

Let L^{xy} ... denote a shorthand notation for $L(P \setminus \{x, y, \ldots\})$. We have the following:

$$L(P) = L^{a} + L^{x}$$
$$= (L^{ax} + L^{ay}) + L^{xa}$$
$$= (L^{axb} + L^{axy}) + L^{ayx} + (L^{xab} + L^{xay})$$

$$= L^{axby} + (L^{axyb} + L^{axyc}) + (L^{ayxb} + L^{ayxc}) + L^{xaby} + (L^{xayb} + L^{xayc})$$

$$= L^{axbyc} + L^{axybc} + (L^{axycb} + L^{axycd}) + L^{ayxbc} + (L^{ayxcb} + L^{ayxcd}) + L^{xabyc}$$

$$+L^{xaybc} + (L^{xaycb} + L^{xaycd})$$

$$= L^{axbycd} + L^{axybcd} + L^{axycbd} + L^{axycdb} + L^{ayxbcd} + L^{ayxcbd} + L^{ayxcdb}$$
$$+L^{xabycd} + L^{xaybcd} + L^{xaycbd} + L^{xaycdb}.$$
$$= 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 = 11$$

Corresponding to the linear extensions $a < x < b < y < c < d$,
$a < x < y < b < c < d, a < x < y < c < b < d, a < x < y < c < d < b$,
$a < y < x < b < c < d, a < y < x < c < b < d, a < y < x < c < d < b$
$x < a < b < y < c < d, x < a < y < b < c < d, x < a < y < c < b < d$,
$x < a < y < c < d < b$, respectively.

$$L(P_{greedy}) = L^{a} + L^{x}$$
$$= L^{ay} + L^{xa}$$
$$= L^{ayx} + L^{xay} + L^{xab}$$
$$= L^{ayxb} + L^{ayxc} + L^{xayc} + L^{xaby}$$
$$= L^{ayxbc} + L^{ayxcb} + L^{xaycb} + L^{xaycd} + L^{xabyc}$$

$$= L^{ayxbcd} + L^{ayxcbd} + L^{xaycbd} + L^{xaycdb} + L^{xabycd}$$
$$= 1 + 1 + 1 + 1 + 1 = 5$$

corresponding to the greedy linear extensions $a < y < x < b < c < d$, $a < y < x < c < b < d, x < a < y < c < b < d$, $x < a < y < c < d < b$, and $x < a < b < y < c < d$, respectively.

$$L(P_{supgreedy}) = L^{a} + L^{x}$$
$$= L^{ay} + L^{xa}$$
$$= L^{ayx} + L^{xay} + L^{xab}$$
$$= L^{ayxb} + L^{ayxc} + L^{xayc} + L^{xaby}$$
$$= L^{ayxbc} + L^{ayxcd} + L^{xaycd} + L^{xabyc}$$
$$= L^{ayxbcd} + L^{ayxcdb} + L^{xaycdb} + L^{xabycd}$$

$=1+1+1+1=4$ corresponding to the super geedy linear extensions

$a < y < x < b < c < d, a < y < x < c < d < b, x < a < y < c < d < b$ and

$x < a < b < y < c < d$, respectively.

6.2 Realizers of a poset

Szpilrajn [33] proved that any order relation is the intersection of its linear extensions. In fact, not very infrequently, the intersection of only a few linear extensions of a given Poset turns out to be sufficient to give rise to the original Poset. That is, for any poset P, there exists a finite set of its linear extensions which realizes P. This leads to the following definition:

If $\mathcal{R}$ is a family of linear extensions of the poset $(P, \leq_p)$ whose intersection is the order relation $\leq_p$,

then $\mathcal{R}$ is called a *realizer* of $\leq_p$. If $\mathcal{R} = \{L_1, L_2, \dots, L_n\}$ is a realizer of the poset $(P, \leq_p)$, then for any $A \subseteq P$, the set $A_{\mathcal{R}} = \{L_1|_A, L_2|_A, \dots, L_n|_A\}$ is a realizer of the subposet $(A, \leq_p)$.

The set of all greedy linear extensions of a poset $(P, \leq_p)$ is a realizer called the *greedy realizer* and the family of greedy realizers is nonempty (Bouchitte et al., 1985). There exists a super greedy realizer for every ordered set (Kierstead et al.,1987).

6.3 Dimension of a poset

The dimension of a poset $(P, \leq_p)$, denoted $\dim(P)$, is defined as the minimum cardinality of a realizer for the poset $(P, \leq_p)$ (Dushnik & Miller, 1941) . In other words, the dimension of a poset is the minimal number of its linear extensions whose intersection is the original poset. That is, $\dim(P)$ is the least positive integer k for which there exists a family $\mathcal{R} = \{L_1, L_2, \dots, L_k\}$ of linear extensions of P such that $\bigcap \mathcal{R} = \bigcap_{i=1}^{k} L_i = P$. It follows from Dilworth's theorem that the dimension of an order never exceeds its width i.e., $\dim(P) \leq w(P)$. Also, it follows from the definition, that the removal of a point from P cannot increase its dimension but it can decrease by atmost one (Hiraguchi, 1951). Thus we have $\dim P \geq dim\{P \backslash \{x\}\} \geq \dim P - 1$ for all $x \in P$. Tree-shaped posets are 2-dimensional. In general, for a poset $(P, \leq_p)$ with $|P| = n$, its upper bound is given by $dimP \leq \left\lceil \frac{n}{2} \right\rceil$. Moreover, if $A \subseteq P$, then $\dim(A, \leq_p) \leq n$.

A Poset $(P, \leq_p)$ (of dimension d) is called *dimension-critical (or d-irreducible)* provided $dim(P \setminus \{x\}) < dimP, \ \forall x \in P$. In other words, P is d-irreducible if it has a dimension $d \geq 2$ and the removal of any element lowers its dimension. The 3-*irreducible* Posets have been characterized, but no characterization is known for the d-*irreducible* posets for $d \geq 4$. However, it is known that for each $d \geq 3$, there exist infinitely many dimension-critical posets of dimension d (Kelley, 1977; Trotter & Moore, 1976). It is shown that the computation of dimension itself is an *NP*-complete problem. In particular, it is polynomial time solvable if a partial order has dimension atmost 2, but the case for having dimension atmost 3 is *NP*-complete (Yannakakis, 1982) . However, whether the jump number is *NP*-complete for the particular class of two dimensional posets is still an open problem (Bouchitte & Habib, 1987).

The notion of greedy dimension of a poset is studied in (Bouchitte et al., 1985). It is observed that the existence of a *greedy realizer* and thus of the *greedy dimension* immediately follows from a result obtained in (El-Zahar & Rival, 1985) that for every incomparable pair (a, b), there exists a greedy linear extension $<_r$ with $a <_r b$.

This is proved by demonstrating algorithmically that such a greedy linear extension exists for every critical pair. Further, in course of studying the relationship between the greedy dimension and the ordinary dimension of a poset, the existence of equality between them for a wide range of posets, including the *N*-free posets, two dimensional posets and distributive lattices has been proved.

Following the definition of $crit(P)$ for a poset $(P, \leq_P)$, a family $\mathcal{R}$ of linear extensions of P is a realizer of P if and only if for every $(x, y) \in crit(P)$, there exists some $L \in \mathcal{R}$ such that $x \geq y$ in L, and hence the dimension of P is just the least integer t for which there exists a family $\mathcal{R}$ of linear extension of P which reverses $crit(P)$ (Trotter, 1991).

Furthermore, if $\dim P = 2$, then every minimal realizer of $(P, \leq_P)$ is greedy.

Since $\mathcal{SG}(P) \subseteq \mathcal{G}(P)$, we note here that

$dim(P) \leq dim_{\mathcal{G}}(P) \leq dim_{\mathcal{SG}}(P)$

holds

(Kierstead & Trotter, 1985).

Besides a wealth of results related to bounds for ordinary and greedy dimensions of a poset, the best possible upper bounds for the super greedy dimension of a poset $(P, \leq_P)$ in terms of $|P - A|$ and width $(P - A)$, where A is a maximal antichain has been proved (Kierstead et al., 1987). Summarily, we have the following:

Removing one point from a poset does not increase any of the parameters: width, height, jump number and dimension. However, it can decrease each of them by atmost one.

If one comparability pair is removed from a poset, its result is a poset in general. However, if only a comparability pair which cannot be recovered by transitivity is removed, the result is still a poset. Thus, removing the comparability $x < y$ results in a poset if and only if y covers x.

Similarly, the addition of one comparability pair does not in general results in a poset. However,if only a comparability which does not *force* other comparabilities is added, the result is again a poset. In other words, the comparability $x < y$ can be added to P with the result being a poset (with exactly one more comparability) if and only if $u < x$ in P implies $u < y$, and $y < v$ in P implies $x < v$. Such a pair (x, y) is called an *nonforcing ordered pair* of P (Rabinovitch & Rival, 1979).

Despite the emergence of consequences that posets exist with bounded height but arbitrary large dimension (Trotter, 1991), numerous significant contributions towards characterizing the dimension parameter of a poset are around (Kelley & Trotter, 1982).

7. Some future directions

(i) In face of the fact that every poset has a greedy optimal linear extension, the characterization for

the existence of non-greedy optimal linear extension of a poset need to be investigated.

(ii) The choice of some useful characterization (say, *stability*, etc.) of a poset interms of the size of its realizers versus the size of the class of its all linear extensions could be investigated further.

(iii) Many nice properties of realizers are known, but how to compute them effectively needs further vindication.

(iv) A number of optimization problems need to be addressed; for example, constructing an efficient algorithm to compute a linear extension that minimizes the number of jumps (Bouchitte & Habib, 1987).

(v) Studies related to various concepts described in this paper on a finite multiset are yet to be conducted (Anderson, 1987; Girish & Sunil, 2009).

(vi) Multiset as a model for *multi-attribute objects* used in discovering intelligent systems, control of Non-linear mechanical systems etc. may get simplified by using topological sorting (Petrovsky, 1997).

(vii) Topological sorting can also be used in discovering computing simulators for biological systems (Krishnamurthy, 2005).

(viii) Last but not the least, some *open* problems, like $1/2 - 2/3$ problem need our attention (Felsner & Trotter, 1993).

References

I. Anderson (1987) , Combinatorics of finite sets, Clarendon Press.

V. Bouchitte, M. Habib & R. Jegou (1985) , On the greedy dimension of a partial Order, order 1, *pp. 219-224.*

V. Bouchitte & M. Habib (1987), NP-Completeness properties about linear extensions, Order 4, pp. 143-154.

Brno Vitezslav Novak (1969), On the well Dimension of ordered sets, *Czechoslovak Math. J. 19 (1*) pp. *1-16.*

Richard A. Brualdi, Hyung Chan Jung & William T. Trotter, Jr. (1992), On poset of all posets on elements, *IMA Preprint series# 976, pp. 1-17.*

O. Cogis & M. Habib (1979), Nombre de sauts et graphes serie- paralleles, *RAIRO Inform. Theorique 13, pp. 3-18.*

R.P. Dilworth (1950), A decomposition theorem for partially ordered sets, *Ann. Math. 51, pp. 161-166.*

R. Ducournau & M. Habib (1987), On some Algorithms for multiple inheritance in Object Oriented Programming, *SEMA-METRA 16-18 rue Barbes, 92126 Montrouge cedex*.

D. Duffus, I. Rival & P.Winkler (1982), Minimizing setups for cycle-free ordered sets, Proc. Amer. Math. Soc., 85, pp. 509-513.

B. Dushnik & E.W. Miller (1941), Partially ordered sets, *Amer. J. Math. 63, pp. 600-610.*

M.H. ELzahar & J.H. Schmerl (1984), On the size of Jump-critical ordered sets,

Order 1,pp. 3- 5.

M.H. El-Zahar & I. Rival (1985), Greedy linear extensions to minimize Jumps,

Disc. Applied Math. 11, pp. 143-156.

S. Felsner & W.T. Trotter(1993), On the fractional dimension of partially ordered sets, *Bell communication research, 445 South street, Morristown, NJ 07962, U.S.A., pp. 1-14.*

K.P. Girish & J.J Sunil, General relations between partially ordered multisets and their chains an antichains, *Math. Commun. 14(2), pp. 193-205.*

T. Hiraguchi (1951), On the dimension of partially ordered sets, *Sci. Rep. Kanazawa Univ. 1, pp. 77-94.*

Jung Hyung Chan (1992), Young's Lattice and Subspace Lattice: Jump number, Greediness,

J. Korean Mathematical Society, 29(2), pp. 423-436.

D. Kelly (1977), The 3-irreducible partially ordered sets, *Canad. J. math. 29 , pp. 367-383*

D.Kelly & W.Trotter (1982), Dimension theory for ordered sets, in Ordered sets

I. Rival, ed., D. Reidel, publishing company, pp. 171-211.

H.A Kierstead & W. Trotter (1985), Inequalities for greedy dimension of ordered sets,

Order 2, pp. 145-164.

H.A. Kierstead, W.T. Trotter & B. Zhou (1987), Representing an Ordered Set as the intersection of Super

Greedy Linear extensions, *D. Reidel Publishing Company, Order 4 , pp. 293-311.*

H.A. Kierstead (1986), NP-completeness results concerning greedy and super greedy linear extensions,

Order 3, pp. 123-134.

Kamil Kloch (2007), Online Dimension of Partially ordered sets,

Reports on Mathematical Logic 42, pp. 101-116.

D.E. Knuth (1973), The Art of computer programming, vol.3 (sorting and searching), Addison-

Wesley, Reading, Mass.

E.V. Krishnamurthy (2005), Rule-based multiset programming Paradigm: Applications to Synthetic

Biology, *Computer Science Laboratory, Austrialian National University, Carberra, ACT 0200, Australia.*

Charalampos Papamanthou (2004), Depth First Search & Directed Acyclic Graphs, Review for the course

Graph algorithms, winter 2004, Department of Computer Science, University of Crete , pp. 1-27.

A.B. Petrovsky (1997), Structuring techniques in multiset spaces. *Springer*, pp. 174-184.

O. Pretzel (1985), Problem presented at Oberwolfach conference on combinatorics and Ordered sets.

I.Rabinovitch & I. Rival (1979), The rank of a distributive lattice, *Disc. Math. 25, pp.275-279.*

I.Rival & N. Zaguia (1986), constructing greedy linear extensions by interchanging chains, *Order 3, pp.*107-121.

K.H. Rosen (1999), Discrete Mathematics and its applications, *McGraw-Hill*.

Bernd S.W. Schröder (2003), Ordered sets, *Birkhauser, Boston*.

Steven Skiena (1997), Topological sort/connectivity, Lecture 16, Department of Computer Science, State University of New York, Stony Brook, N.Y .

E.Szpilrajn (1930), Sur I' extension de I' ordre partiel, *Fund. Math. 16, pp. 386-389.*

William T. Trotter & J.I. Moore (1976), Characterization problems for graphs of partially ordered sets, lattices, and families of sets, *Disc. Math. 15, pp. 71-82.*

William T. Trotter, J.I. Moore & D.P. Sumner (1976), The dimension of a comparability graph, *Proc. Amer. Math. Soc. 60, pp. 35-38.*

William T. Trotter (1991), RAMSEY Theory and Partially ordered sets, pp. *1-11*.

M. Yannakakis (1982),The complexity of the partial order problem, *SIAM J. Alg. Disc. Math 3, pp. 351-358.*

J. Valdes, R. E.Tarjan & E. L. Lawler (1982), The recognition of series parallel digraphs, SIAM J. Computing 11 (2), pp. 298–313.

DEVELOPMENT OF A COMPREHENSIVE CITY STRUCTURE DATA BASE FOR THE PLACEMENT OF BASE STATIONS: A CASE STUDY OF MINNA, NIGERIA

Caroline O Alenoghena[1],

Joy O Emagbetere[2]

&

Musa A Aibinu[3]

[1]Federal University of Technology, Minna, Nigeria.
(alenca2@yahoo.com)
[2]University of Benin, Nigeria.
(joyokumo@yahoo.com)
[3]Federal University of Technology, Minna, Nigeria.
(maibinu@gmail.com)

***Abstract*:** In this paper, a spatial database called "City Structure Database System" (CSDS) collection process is developed. This database is to represents situations in real-world by maps and digital data. With the use of the database, radio planners can access and get hold of relevant map information of a city easily. An effective data collection process is developed for the acquisition and collation of spatial data of the city. This database is required by an intelligent decision making process of base station placement. Experimental results of implementation in Minna town Nigeria, reveal the effectiveness of the method proposed in this paper.

Key words: Base station, City data base, GPS, Spatial data, Matlab programme,

1.0 Introduction

The explosion and rapid demand for wireless communication services has resulted in an increase in the number of base stations in our environment. While research is still on going to ascertain the medical impact of the base station within our living quarters efforts are still being made to make their presence less intimidating and acceptable. Many human activities require geographical information and the placement of base station is one of such. In wireless communication networks the choice for location for the network infrastructure is a tedious exercise usually carried out during the planning process. The availability of a comprehensive city data base proposed will go a long way to providing a more efficient

and effective infrastructure deployment process. The placement of base stations without adequate spatial and thematic information can be said to be optimally implemented. These lapses can be attributed to the non availability of such information and the dependence on their acquisition by government agency only. In other to locate a facility there is the need for comprehensive information of existing and potential structures which will be of concern to the infrastructure. In this paper we develop a user friendly process for the acquisition and collation of geographical data to be used in an intelligent decision making process for base station placement.

One of the key infrastructures of a mobile communication network is the base station. The base station is the facility placed to provide radio coverage for a given area referred to as cell. The optimally placement of base station has become a concern with the many constraints and contradicting factors to overcome [Alenoghena, 2012]. Some of these constraints include environmental aesthetics, safety, topography, demographic traffic distribution of users, interference between network cost, etc. The base station placement algorithm is a heuristic one genetic algorithm which will be searching a wide space.

In the remaining part of this paper, we discussed the need and significance of the study. The database organization and implementation is thereafter presented. Some data collected are presented before conclusions.

2.0 Need for the Study

In other to place base stations, site selection has to be carried out. The present manual exercise involved in site selection and survey process has been described as rigorous, inefficient and could be prone to errors as they are based on personnel experience and judgment [Job, 2008] This work is aimed at creating a detailed database of site information that will reduce the rigor and inefficiency prone in the usual way site surveys and selection process is done. The non availability of a Geographical information system database in most developing countries necessitated this research aimed at providing a template for digital data acquisition, storing and retrieval system that is user friendly and effective for radio planners. Information in the database can be used by other agencies, government and for developmental purposes.

Most developed countries have their cities/municipal supported by geobased online applications and being reffered to as ‘smart cities’. The City of Edinburgh for instance has developed a very modern and impressive internet presence with its services ranging from online maps, historic maps, catchment areas of school with detailed statistic. Information provided include road works, and traffic flow situation in

real-time [smartcities]. This proposed database will be of immense benefit for non smart cities.

2.1 Significance of this Study

The base station (BTS) is the most important section in the Global system for mobile communication (GSM) services as it is the link to the final consumer. Its proper location ensures quality service and interactions with the public. The agitation for more base stations and issues of non compliance of most GSM base station in the country to stipulated safety requirement by regulatory agencies in Nigeria necessitated this research work. In cellular telecommunication system deciding where to place the base station is a very important issue during the process of cell planning [Ajay, 2004]. The BTS for a given geographical area has to be well positioned for maximum coverage and minimal interference which are the indexes for measuring the quality of any mobile service [Kia et al, 1998]. However, if coverage is the only issue of contention in cell planning then the BTS would have been placed anywhere. However, if it were so, the issue of environmental safety violation as well as the quality of services provided would be compromised since the signal to interference ratio may be minimized. This is the actual scenario in the Nigeria mobile cellular communication industry.

This fact was buttressed in the research carried out by Okonighene [2010], that in Nigeria, 98% of (GSM) base stations (cell sites) are sited within 20 meters from residences, offices, schools, business buildings, petrol stations and public arenas. According to the environmental regulatory body [nesrea.org] this is a violation on environmental safety. The presence of a BTS is always felt by the surrounding structures, human and material, irrespective of the fact that the BTS may stand alone on a space with the minimum space as specified by regulations. The need to handle public concerns on health effect of the BTS has been discussed by several researchers. [Julie, 2005; Biebuma, 2011; Danielle, 2005]. In recent times, the Nigerian environmental regulatory body (NESREA) have had course to shut down some base station that pose threat to public safety [guardiannews.com].

This research considered the provision of a digital city data base will aid the radio planning engineer to make a cost resourceful site selection. The data base together with the other radio network planning tools will provide the radio engineer an informed knowledge of the territory to be covered by a proposed base station. This will go a long way in ensuring the selection of suitable sites for BTS placements, taking environmental safety requirement into consideration. The

spatial data base generated will also serve for other facility placement decisions in the state.

3.0 Data base Organization

The city database will hold records on the land surface of the city identified by its geographic location. The city is divided into plots/units of 50m x 50m with the key field or field name being the Lat/Log as it is unique for every point in the universe. The city database has two other fields which are the elevation from sea level in meters and the class or nature of existing infrastructures if any. Table below shows the different fields for the proposed database.

Table 1.0: Data base fields.

Longitude	Latitude	Elevation	Class

The class refers to the type of feature on ground a specific location, for a comprehensive database the following range have been identified as class Residential, Commercial, Industrial, Civic/ cultural, Health/clinics, High ways, Vegetation and Open space. A ground survey has to be carried out to determination the class of each entry. The city database is organized following the tabular attribute format of data base management system. The database consist of a two dimensional table with the column indicating the attribute and the rows the records. The database has three main attribute which are the coordinates (latitude and Longitude) the elevation from sea level in meters and the class. The Class attribute is linked to a relational table where the class categories are defined. Table 2.0 shows the class categories and assigned class code.

Table 2.0 shows the different class categories

Class code	Class Category
1	Residential
2	Commercial
3	Industrial
4	Civic/cultural
5	Educational
6	Health /clinics
7	Highways
8	Vegetation

This class code is the major link between the two tables 1.0 and 2.0, developed following the relational DBMS model for storing attributes [Urisa, 1999]. The common and unique identifying attribute in this case is the 'class code'.

4.0 Implementation:

The implementation of the database collection is thorough but with proper organization the task was achieved in a short time. The entire city/territory was divided into pixels of units which are then subdivided into smaller units for better organization. The data upload was carried out using a programme written in Matlab. The entire city is sectioned into different regions with each regions cut out into several plots as depicted in the Minna city map shown in fig 1.0

Fig 1.0: Minna City Map showing area demarcation

The upload programme is so organized for adequate continuation and saving of data collected. The spatial data for small regions implemented were taken using a Germin handheld GPS device with a precision of 2 meters. The flow chart for the data upload is as shown in figure 2.0

4.1 Programme features

Unique features have been included in this programe to allow for a comprehensive data collection.

These include the data input template with the various field provided. A map data upload indicator that allows for indicating in the form of a black square the region within the uploaded map whose data has been captured.

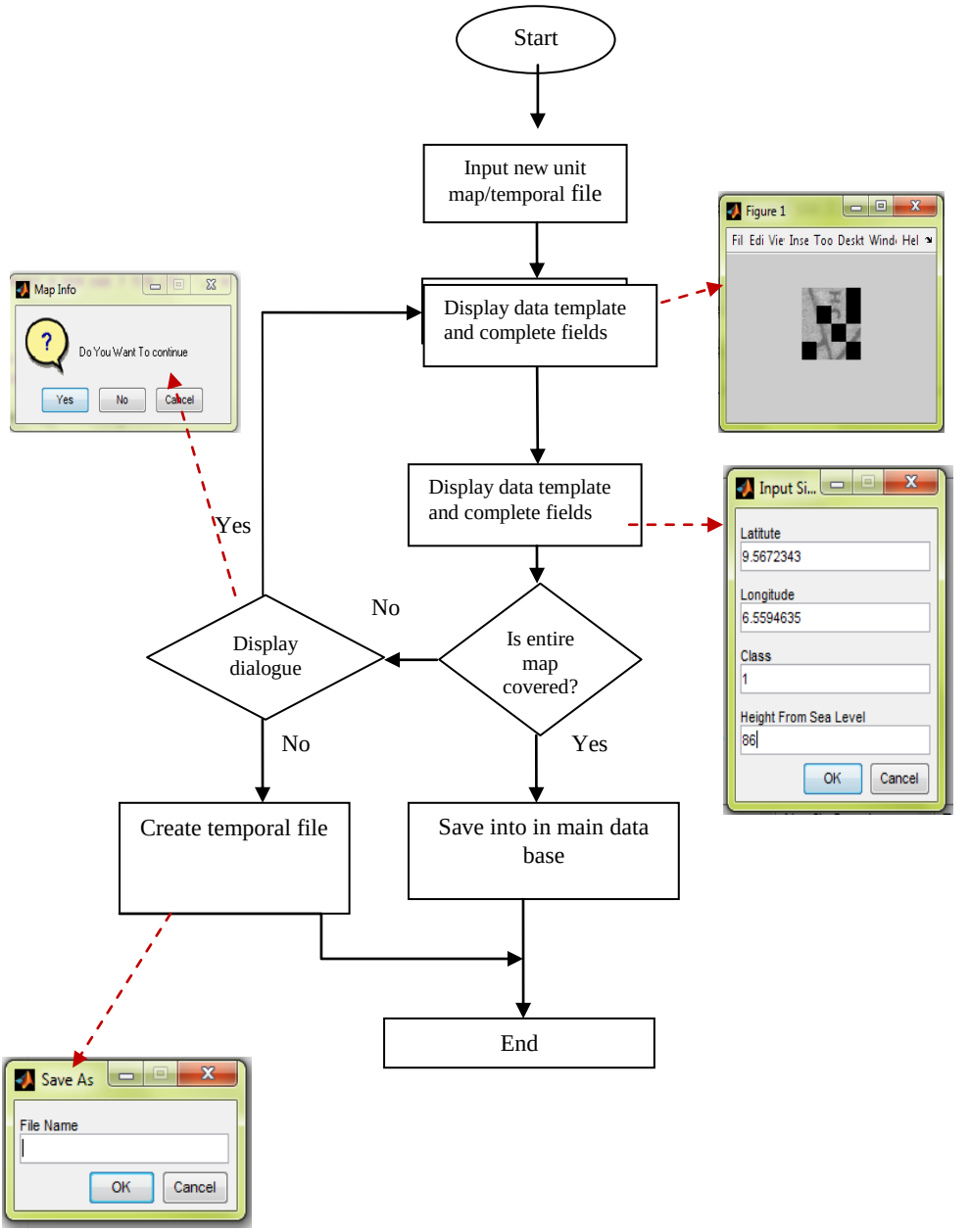


Fig 2.0 Flowchart for data upload in Matlab environment

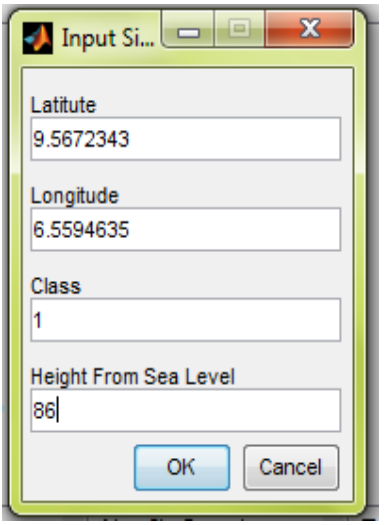


Fig 3.0 Data input template

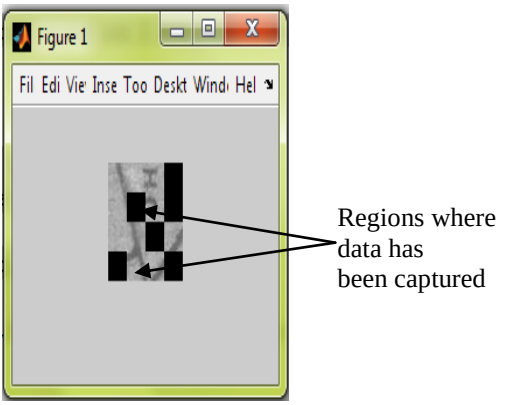


Fig 4.0 Map data upload indicator

The data saving and continuation process is automated in line with modern GIS database repository format for storing databases. Figures 5.0 a, b, and c show the graphical user interfaces for saving, and retrieving of the spatial data.

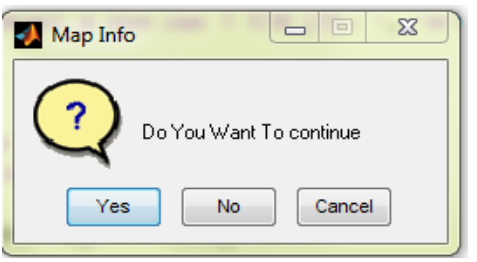


Fig 5.0 (a) Inquiry for contuation of data up load

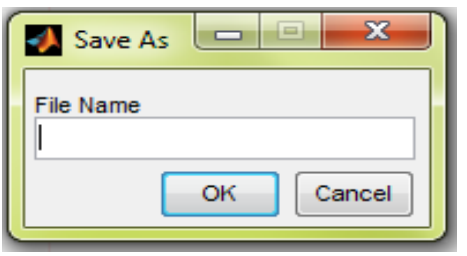


Fig 5.0 (b) Creating a tempoary file name

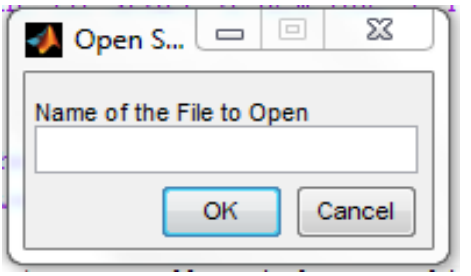


Fig 5.0 (c) Opening a previously saved file

4.2 Area of Implementation

The developed template and image map building was implemented in Minna, the capital of Niger state, Nigeria. Minna is located at approximately between latitude 09^0 25'N and $09^0$45'N; and longitude 06^0 15'E and 06^0 35' E with an average elevation/altitude of 299 meters [Ojigi 2012]. The terrain in Minna can best be described as asymmetrical with surrounding hills. The inhabited regions in the town however are flat, with very scanty habitation in the hilly regions. Sample of some data taken in Minna is given in table 3

Table 3: Raw field data of data in Kpakungu Minna

Longitude	Latitude	Elevation	Class Code
6.533167	9.597778	239	2
6.532694	9.597639	240	1
6.532361	9.597444	240	2
6.531805	9.597306	235	4
6.531333	9.597083	236	5
6.531833	9.597000	234	2
6.531055	9.596278	239	2
6.531333	9.596778	234	2
6.531138	9.596806	235	6
6.530666	9.596833	237	1
6.530361	9.596583	234	1
6.529305	9.596361	232	1
6.529472	9.595833	233	1
6.530722	9.596889	233	1
6.529583	9.595333	232	1
6.529638	9.594583	231	1

6.530083	9.594556	231	1
6.530416	9.595000	230	1
6.530888	9.595500	232	5
6.531083	9.595833	232	6
6.531500	9.596333	230	1
6.531777	9.596083	230	1
6.531944	9.596194	230	1
6.532138	9.596583	231	1
6.532694	9.596722	230	1
6.533305	9.596861	231	1
6.516888	9.596250	232	1
6.517166	9.596889	234	1

Conclusion

Spatial data collection is a GIS base activity with accuracy achieved through the use of high precision GPS facilities. In situations where this information is not available as is the case in Minna, Nigeria, facility locations such as the base station placement has been a challenging task. The easy access and retrieval of this information are very necessary. In this paper a comprehensive data collation process have been presented, with fields relevant to the information required by radio planners. The generated database will be used to test an ongoing work on base station placement using genetic algorithm. The accuracy of data collected will be verified in future work with the development of a digital map displaying ground features.

References

Ajay M. (2004) "Fundamentals of Cellular Network Planning and Optimization" John Wiley & Sons Ltd, England (2004) pgs 21-31

Alenoghena C. O, and Emagbetere J. O (2012) "Base Station Placement Challenges in Cellular Networks: The Nigerian Experience. "Proceedings of 4th IEEE International Conference on Adaptive Science & Technology (ICAST) Ghana 25-27th Oct 2012. (pages 7-11)

Andrew B, Benjamin H and Iain L. (2013) "mHealth Possibilities in a Changing World. Distribution of Global Cell Phone Subscriptions" Journal of mobile technology in medicine Electronic ISSN 1839-7808. DOI:10.7309/jmtm.78.

Biebuma J. J., Esekhaigbe E. (2011) "Assessment of radiation hazards from mobile phones

and gsm base stations" Journal of Innovative Research in Engineering and Science 2(1), March, 2011. ISSN : 2141-8225

Danielle S. and Olivier B. (2005) Regulating the risks of mobile phone base stations, a comparative study in 5 European countries Proceedings of International Workshop on Base Stations and Wireless Networks: Exposures and Health Consequences Switzerland, Geneva June 15-16, 2005 pgs 113-126

Guardian Newspaper [2012] "NESREA-NCC Continue Fight over Base Stations" Tuesday 08 May 2012 http://www.ngrguardiannews.com

Julie B, Lada T, Richard S, Vicky S and Jane V, (2005) "Understanding Public Responses to Precautionary Action and Advice" proceedings of International conference on base station and wireless network: Exposures and health consequences, Switzerland Geneva June 15-16th 2005

Job M., Anish K., Ben V. W., (2008) "Optimization of Antenna Placement in 3G Networks using Genetic Algorithms" *Proc IEEE Third International Conference on Broadband communication, Information Technology and Biomedical Applications pg. 30 – 37*

Kai L., Erkki L.' and Jaakko L. (1998) "Radio Coverage Optimization with Genetic Algorithms" The 9th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication vol.1 pg 318-312.

Kindra S, Thomas W, and Keith W(2006) "Comparing GPS Receivers: A Field Study" Journal of Urban and Regional Information System Association (URISA) Vol. 18, No. 2 pgs 19-23.

MDZmultimedia (Nig.) Ltd. June 2005, "Implementing GIS & creating spatial data" www.mdzmultimedia.com

Ojigi L. M (2012) "An Evaluation of the Efficiency of the Land Use Act 1978 of the Federal Republic of Nigeria and its Implications in Minna and Environs" Pro Poor Land Management, 5577 FIG Working Week 2012, Rome Italy, 6-10 May 2012.

Okonigene, R.E.; (2010) Siting of GSM base stations antenna and its Health Consequences" Information Technology: New Generations (ITNG), IEEE seventh International Conference on Information Technology. Pgs 613 – 618.

Schonhofer M. (2005) "Optimizing Base station Sitting in Umts Networks" Diploma thesis Institute of Communications

and Radio-Frequency Engineering Vienna University Of Technology

URISA (1999) 'Gis Database Concepts" Urban and regional Information system association (URISA), Quick lecture series, ISBN #: 0-916848-09-4 USA. Available on line at www.urisa.org

http:// www.nesrea.org/ - The National Environmental Standards and Regulations Enforcement Agency (NESREA) official website. Assessed on 20th September, 2012

http://www.worldatlas.com/aatlas/imageg.htm ***Latitude and Longitude*** - World Map with ***Latitude and ...*** - World Atlas. Assessed on 5th April 2013

IMPACT OF CODING PHASE ON OBJECT ORIENTED SOFTWARE TESTING

Sanjeev Patwa[1]

&

Anil Kumar Malviya[2]

[1]FASC, Mody University of Science and Technology,
Lakshmangarh, Sikar, Rajasthan,
[2]KNIT, Govt. Engg. College, Sultanpur, U.P.
[1]sanjeevpatwa@yahoo.com, [2]anilkmalviya@yahoo.com

Abstract: The paper demonstrate the findings of empirical research from 23 software development companies to identify the factors of coding phase which effects the testing of Object Oriented (OO) software. Six potential factors of coding phase have been identified. The study uses a relative weight method and ANOVA test to analyse these factors and identify the place of these factors according to their impact on Object Oriented software testing. The survey focuses on the crucial participants like programmers and testers who highly involve in coding phase.

Keywords: ANOVA, Factors of Testing, Object Oriented Software, Relative weight method.

I. Introduction

The software development process consists of five phases: analysis, design, coding, testing, and operation (Lee et.al., 1999).According to (Beizer, 1990), testing is considered as one of the most costly development processes, sometimes exceeding fifty per cent of total development costs. In Object-Oriented systems the majority of development is concerned with analysis, design, and coding phases. Each phase of development has its own importance in testing. Testing is a continuous process involves in each phase of software development but software testing has not kept pace with the advances in OO system programming. Some OO systems do not conform to traditional testing definitions or techniques. There are several factors in each phase which affect the software testing. In this paper we are considering only coding phase and related factors.

Although the question of what are the factors which affect the testing techniques is a question which testers face every time they have to test a system. In general there are several testing techniques are exists,

but it is unfortunate that some are never considered for use, and others are used over again in different projects without even observing, whether or not they were really useful. There are basically two reasons why developers do not make good choices:

- The information about testing techniques are distributed across academic tools like journals, books, articles etc. and technical people like developers and testers rarely read or study the literature.
- They have fewer tendencies to share the technical experience with others of using testing techniques in different projects. This means that they miss out the chance of learning about the experiences of others.

The problem we emphasis here is to identify relevant factors and their places with coding phase, which affect the testing techniques in OO software and to find the opinion of programmer and tester on significance of selected factors for, object oriented software testing. The aim of solving this problem is to help testers to find factors which affect the testing of project in coding phase and select the suitable testing technique

II. Related Work

The problem of identifying and analyzing the factors which affect the testing in coding phase has not yet been specifically studied in the testing area, there have been some attempts which focuses on comparing the testing techniques. There are not solutions for the problem on which we are emphasizing here, although the knowledge they input can be used in analyzing the problem that we propose.

A main problem with testing object-oriented systems is that standard testing methodologies may not be useful. Smith and Robson (Smith & Robson, 1994) say that current IEEE testing definitions and guidelines cannot be applied blindly to OO testing. Harold (Harold et al., 1992) present a technique that takes advantage of the hierarchical nature of classes, utilizing information from the super classes to test related groups of classes. Parrish (Parrish et al., 1993) present a technique for testing OO systems that is based entirely on generating test cases from the class implementation. McGregor (McGregor and Korson, 1994) discussed a high-level view of testing OO systems within the entire software development cycle. Research on software process innovation found that factors, such as organization size, technological diversity, professionalism, and IS slack are related to the adoption process (Grover et.al, 1997). Characteristics of a software organization, such as IT infrastructure, staff, managers, and control systems, have an effect on identifying, adopting, testing and implementing an innovation (Tornatzky & Fleischer, 1990).These attributes and organizational characteristics

comprise the basis for our research model and hypotheses.

In previous some years, many software testing techniques has been proposed for achieving software reliability. However in early stages of software development mostly all the factors which affect the testing are ignored for simplicity reasons. But in order to improve the understanding and simulation of a complex process such as the software coding process, factors like Programmer/Tester skill, Programmer/Tester organization (the percentage of high-quality programmers), Development team size, Program workload (stress), Domain knowledge, Human nature (mistake and work omission) need to be considered and incorporated into selecting the software testing techniques.

To our knowledge very less research has been conducted in the area of finding the factors and analyse their impact on testing of coding phase, and how people consider these factors in enhancing the software testing process.

III. Research Goals

In this study, we performed a survey and obtained opinion from programmers and testers who participated in the software development. In this research we identified six factors which involve in the coding phase of the software development process.

The paper emphasis on the hypothesis that:

H0: All the selected factors have significant impact on testing of object oriented software in coding phase.

In hypothesis test the confidence-level is consider at 95% (at 5% level of significance) that means if the value of p is less than 0.05 ($p < 0.05$) we reject the null hypothesis.

Software development is a very complex process which involves several factors. These factors vary from project to project. In our modern society, object oriented software has become a very critical component in all kinds of systems and software.

The factors considered in this study are Programmer/Tester Skill, Programmer/Tester Organization (the percentage of high-quality programmers), Development Team Size, Program Workload (stress), Domain Knowledge and Human Nature (mistake and work omission).Some of these factors also considered by Zhang (Zhang & Pham, 2000) for measurement of software reliability. The description of factors is as follows:

1. Programmer/Tester Skill: Skill can be defined as the average number of years of programming experience of programmers. This can be calculated as the ratio between total year of experience of all programmers/tester and total no of programmers/testers in organization.

2. Programmer/Tester Organization: Programmer /Tester organization (PO) is defined as the percentage of high-quality programmers. PO is computed as follows: PO = n_h /n ; where n_h is the number of programmers, whose programming experience is more than six years, and n be the total number of programmers in organization. As PO is high, we can select better testing technique [ix].
3. Development Team Size: This factor identifies that quality of project would improve if the size of team will be large or the quality will improve with the less but experienced development team.
4. Program Workload: During the software development, stress factors in terms of "work contents" such as schedule pressure and too much work are the major factors. This factor may affects in selecting the Testing technique.
5. Domain Knowledge: Domain knowledge refers to the programmer's and tester's knowledge of the input space and output result. Insufficient knowledge may cause problems in coding and testing procedures.
6. Human Nature: This refers to the tester and developers characteristics, including the ability to avoid the making of working mistakes, careless work omission and selecting the testing technique.

The results of this study may be utilized in selecting the testing techniques in coding phase by incorporating significant factors. This study aims to present the rank of factors according to their significance in OO software testing.

IV. Statistical Methodology

Software testing in coding phase, in the context of analysing factors is a new area of research; this study is exploratory in nature yet specific in view of the conceptual models. Field examination through questionnaire and study was chosen as the overall design approach. The factors are described in table I.

Table I. Factors of Coding Phase

Coding
Programmer/Tester Skill
Programmer/Tester Organization
Development Team Size
Program Workload (stress)

Domain Knowledge
Human Nature (mistake and work omission)

We utilized Relative weight method to analyze the ranking of factors and for hypothesis testing parametric test ANOVA (Analysis of Variance) has been used.

Places and determining the relative weights of factors based solely on the participants opinions as reflected in the questionnaire. Under this methodology, we treat every single participant equally without considering his/her background information.

V. Data Collection

The questionnaire focused on factors which affect the testing in coding phase and try to examine the view of participants.

Data were collected using a formal survey questionnaire given directly to the software developers/programmers and testers in 23 Indian software development and testing organizations including focusqa.com, pure testing software Pvt. Ltd., TCS, Metacube Systems, etc. Demographic data on the participants are summarized in Table II.

Questionnaire used a 5-point scale to identify the degree to which each factor (the independent variables) has significant influence on software testing in coding phase. In the questionnaire, "1" indicates "not significant" and "5" stands for "most significant".

Table II. Demographic data of survey participants

Personal/Demographic factor	Mean score	Sample size
1. Current job Position (distribution of the survey participants)	Programmer: 72.61% (122)	168
	Tester: 27.38% (46)	
2. Number of experience (years)	4.67=5 Years	156
3. % of people agreed on significance of factors	82.85%	168

VI. Hypothesis Test and Analysis

A. Relative weight method

First, the relative weight method was used to obtain the final places for the factors. Let rij be the original ranking of the i th factor on the j th survey. We first normalize these rij's such that

$$wij = \frac{rij}{\sum_{i=1}^{n} rij} \qquad (1)$$

Where n is the number of factors on the jth survey. Therefore $\sum_{i=1}^{n} wij$=1 for all j.

Different people may give different original ranking and some of them may give higher scores for all factors. By normalizing the original ranking scores using Eq. (1), one can get rid of this bias. We then average these wij's to obtain the final weight for the ith factor such that

$$wi* = \frac{\sum_{j=1}^{l} wij}{l} \qquad (2)$$

Where l is the number of surveys used in this method. Based on these relative weights, we could obtain the final weight for each factor.

From the results by the relative weight method places of factors is given in Table III. The column named Normalized Priorities gives the contribution of each factor. For example, Programmer/Tester Skill contributes approximately 4.4% (its relative weight. 0.0446987). Higher priority value indicates a higher place. Since lower class rank implies decrease in magnitude of relative importance, software programmer and tester should then pay more attention to the factors with high places. The final priority information can then be used to guide the Object Oriented software testing process in coding phase of different applications.

Table III. Final ranking based on relative weight method

Rank	Rank factor	Factor Name	Normalized Priorities
1	F1	Programmer/Tester Skill	0.0446987
2	F5	Domain Knowledge	0.0430985
3	F6	Human Nature (mistake and work omission)	0.0369283
4	F2	Programmer/Tester Organization	0.0348077
5	F4	Program Workload (stress)	0.0339011
6	F3	Development Team Size	0.0226598

From the demographic data in table II it is observed that 83% participants agreed on the significance of impact of selected factors on testing techniques. Table III indicates the top two factors of coding phase are Programmer/Tester Skill and Domain Knowledge.

B. ANOVA test

We conduct ANOVA test for each factor and analyse the statistics as, the testing hypothesis is accepted if the p-value is more than 0.05 at 95% level of confidence, otherwise the testing hypothesis is not accepted (rejected). Resultant tables and descriptions are as follows:

Programmer/Tester skill (F1):

Table IV ANOVA test for Programmer/Tester Skill

		Sum of Squares	df	Mean Square	F	Sig.
F1	Between Groups	7.723	2	3.861	2.675	0.072
	Within Groups	238.182	165	1.444		
	Total	245.905	167			

Table IV evident that the computed value of F-statistics is 2.675 which is less than tabular value of F statistics (3.00) therefore the Null hypothesis is accepted at 5% level of significance for this factor. Subsequently same result can be established with respect to p-value which is 0.075. Therefore, it is concluded that the Programmer/Tester Skill significantly affect the testing in coding phase.

Programmer/Tester organization (F2):

Table V –ANOVA test for Programmer/Tester Organization

		Sum of Squares	df	Mean Square	F	Sig.
F2	Between Groups	15.911	2	7.955	5.064	0.007
	Within Groups	259.208	165	1.571		
	Total	275.119	167			

It is evident from table V that the computed value of F-statistics is 5.064 which is more than tabular value of F statistics (3.00) therefore the hypothesis is rejected at 5% level of significance. Subsequently same result can be established with respect to p-value which is 0.007. Therefore, it is concluded that the Programmer/Tester Organization significantly affect the testing in coding phase.

Development Team Size (F3)

Table VI ANOVA test for Development Team Size

		Sum of Squares	df	Mean Square	F	Sig.
F3	Between	24.407	2	12.204	5.723	0.004

	Groups					
	Within Groups	351.872	165	2.133		
	Total	376.28	167			

Table VI evident that the computed value of F-statistics is 5.723 which is more than tabular value of F statistics (3.00) therefore the hypothesis is rejected at 5% level of significance. Subsequently same result can be established with respect to p-value which is 0.004.
Therefore, it is concluded that the Development Team Size less significantly affect the testing in coding phase.

Program workload (stress) (F4)

Table VII ANOVA test for Program Workload (Stress)

		Sum of Squares	df	Mean Square	F	Sig.
F4	Between Groups	22.628	2	11.314	10.182	0.00
	Within Groups	183.348	165	1.111		
	Total	205.976	167			

It is evident from table VII that computed value of F-statistics is 10.182 which is more than tabular value of F statistics (3.00) therefore the hypothesis is rejected at 5% level of significance. Subsequently same result can be established with respect to p-value which is 0.00.
Therefore, it is concluded that the Program Workload less significantly affect the testing in coding phase.

Domain Knowledge (F5)

Table VIII –ANOVA test for Domain Knowledge

		Sum of Squares	df	Mean Square	F	Sig.
F5	Between Groups	7.215	2	3.608	1.841	0.162
	Within Groups	323.261	165	1.959		
	Total	330.476	167			

Table VIII represent that the computed value of F-statistics is 1.841 which is less than tabular value of F statistics (3.00) therefore the hypothesis is accepted

at 5% level of significance. The same result can be established with respect to p-value which is 0.162.
Therefore, it is concluded that the Domain Knowledge significantly affect the testing in coding phase

Human Nature (mistake and work omission) (F6)

Table IX ANOVA test for Human Nature

		Sum of Squares	df	Mean Square	F	Sig.
F6	Between Groups	22.203	2	11.101	6.929	0.001
	Within Groups	264.369	165	1.602		
	Total	286.571	167			

It is evident from table IX that the computed value of F-statistics is 6.929 which is more than tabular value of F statistics (3.00) therefore the hypothesis is rejected at 5% level of significance. Subsequently same result can be established with respect to p-value which is 0.001.

Therefore, it is concluded that the Human Nature less significantly affect the testing in coding phase.

From above discussion it is evident that hypothesis has been accepted for two factors out of six. The two factors Programmer/Tester Skill and Domain Knowledge also has highest ranking according to relative weight method in table III. Other four factors do not accept the null hypothesis means they are less significant for testing in coding phase this result also verified the ranking of factors indicated in table III.

VII. Conclusion and Future Work

This paper shares out with the factors involved in the testing of coding phase in Object Oriented Software development process. A study was performed to collect the data. The relative places of the factors have been provided in terms of the significance of their impact on software testing. Developers and testers can check the list and find out the most significant ones for their projects. From the opinion of our study participants, we can see that most of the people (82.85%) of the respondents agreed on the fact that selected factors influence on software testing.

However survey result shows that the factors, Programmer/Tester Skill (F1) and Domain Knowledge (F5) significantly affect the testing of Object Oriented software in coding phase while other factors which less significantly affect the testing are as

follows: Programmer/Tester organization (F2),

Development Team Size (F3), Program workload (stress) (F4) and Human Nature (mistake and work omission) (F6).

The findings, however, are based on the group of people who participated in the questionnaire. Caution need to be taken when applying these results in other applications. This study provides a basis for many different directions for further research one of them is to introduce more factors in the questionnaire.

References

Beizer, B. (1990). Software Testing Techniques, Second Edition, Van Nostrand Reinhold Company Limited.

Grover, V., Fiedler, K. & Teng, J. (1997). Empirical evidence on Swanson's tri-core model of information systems innovation. Information Systems Research, 89(3), 273-287.

Harold, M., McGregor, J., & Fitzpatrick, K. (1992). Incremental Testing of

Object-Oriented Class Structures. Proceeding of the International Conference

on Software Engineering, 68 – 79.

Lee, M., Pham, H., & Zhang, S. (1999).A methodology of priority settings its application on software development. European Journal of Operational Research.

McGregor, J., & Korson, T. (1994). Integrated Object-Oriented Testing an Development Processes. Communications of the ACM, 59-77.

Parrish, A., Borie, R., & Cordes, D. (1993). Automated Flow Graph-Based

Testing of Object-Oriented Software Modules. The Journal of Systems and Software, 95–109.

Smith, M., & Robson, D. (1994). A Framework for Testing Object- Oriented

programs. Journal of Object-Oriented Programming, 45 – 53.

Tornatzky, L.G. & Fleischer, M.(1990). The Processes of Technological

Innovation, Lexington, MA: Lexington Books.

Zhang, X., & Pham, H.(2000). An analysis of factors affecting software reliability. The Journal of Systems and Software, Elsevier, 50, .43-56.

Authors Biography:

Mr. Sanjeev Patwa is Lecturer in Computer Science at FASC, Mody University of Science and Technology, (MUST), Lakshmangarh (Sikar) .He has long teaching experience. He received his B.Sc. (PCM) from M.L.Sukhadia University Udaipur , MSc (C.S.) & M.C.A. from JRN Deemed University, Udaipur. He also received "B" Level certificate from DOEACC Society, New Delhi. He has also done MTech (IT) from KSOU, Maysore. He has published 12 papers in National & International Journals and 10 papers are in International/National conferences and seminars. He is also author of 4 computer science books. His research interest is Software Engineering. He is Life Member of Rajasthan Science Congress.

Dr. Anil Kumar Malviya is an Associate Professor in Computer Science & Engineering Department at Kamla Nehru Institute of Technology, (KNIT), Sultanpur. He received his B.Sc. & M.Sc. both in Computer Science from Banaras Hindu University, Varanasi respectively in 1991 and 1993 and Ph.D. degree in Computer Science from Dr. B.R. Ambedkar University, Agra in 2006. He is Life Member of CSI, India. He has published over 40 papers in National & International Journals and 5 papers are in International/National conferences and seminars. His research interests are Data mining and Software engineering, Cryptography & Network Security.

EVOLVING PREDICTOR VARIABLE ESTIMATION MODEL FOR WEB ENGINEERING PROJECTS

Sanjay Kumar Srivastava,

Poonam Prasad &

S.P. Varma

Department of Mathematics and Computer Science,B.R.A. Bihar University, Muzaffarpur, India
srivastavasanjaykumar@yahoo.co.in,
poonamprasadmca@yahoo.co.in,
spvarma2shailesh@gmail.com

Abstract: The field of web cost estimation is an important area which has not received much attention and has been the reason of failure of a number of web projects. Early web cost estimation can save disastrous situation. In this paper a review of previous work done in the field of web estimation like Size measures of Cowderoy, Mendes et al., Rollo and Cleary has been made. It further narrates Mendes Web Cost estimation Model, Tukutuku Project and derivation of effort equation. The paper investigates the applicability verification survey of web projects developed by Indian companies. The verification results being favorable Evolving Predictor Variable Estimation Model for Web Engineering Projects has been framed on the basis of curve fitting by sums of exponentials using Froberg's method and Moore's method. The model takes into the strength of development team as predictor variable and gives the gross effort in hours worked by taking into account the additional issues of contingencies, risk management and profitability issues which were neglected by Mendes et al. while creating their Early Web Cost Estimation Model. It further calculates the Current Sale Price of the web based project based on the development team in Indian Rupees (INR) at the 2014 price level. The early web effort and cost prediction has thus become more accurate and shall be advantageous at the project enquiry and bidding stage.

Keywords/Index Terms*:* Size Measures, Metrics, Early Web Cost Estimation.

1. Introduction

The Internet, it has been said, is the greatest invention of the twentieth century. Internet is an acronym for internetwork and is in fact a network of computer networks that may be dissimilar and are joined together by means of gateways that handle data transfer and conversion of messages from the sending networks' protocol to those of the receiving ne

World Wide Web (www) or simply the Web is the total set of interlinked hypertext documents residing on the Hyper Text Transfer Protocol (HTTP) servers around the world of Internet. The documents on the Web, called Web pages, are written in Hyper Text Markup Language (HTML) and are

identified by Uniform Resource Locaters (URLs).

Web Engineering can be defined as the application of the principles of mathematics and science in order to create and utilize the web pages efficiently and effectively. Web Engineering Projects are concerned with the Web Applications (web apps) development, economically, timely and with least efforts as far as possible.

The software industry is facing greatly enhanced competitiveness due to the emergence of market leaders from the third world economies like India, Korea, Mexico, Malaysia etc. and China is the latest addition to the list. This has added fuel to the fire in the global recession that has severely affected the IT sector worldwide. In fact the need of the hour is to evolve cutting edge estimation techniques which deliver accurate estimates of size, effort, schedule and cost of software development projects. There has been rising trends towards the development of the network compatible web engineering projects catering to the needs of the multinational corporations in this era of business globalization. The time has come for an insight and in depth study and analysis into the factors affecting the web based projects' estimation [1]. The emergence of web as a delivery environment has catapulted both the commercial and educational web application development. Although a variety of development tools are available to the web developers yet the industry lacks a uniform and standardized development technology. The scarcity of data sets relating to the web projects development history and the lackadaisical approach of the web development industry has been a major repulsion for researchers in this area. The problem is further compounded in India by fierce competition in the software industry where the web development companies has been often trying every trick and resorting to every technique. They view with suspicion when asked for their past development data sets relating to records of project development effort, cost and time schedule etc. mistaking researchers with persons spying on behalf of rival companies eliciting information to be misused against them in forthcoming tenders and rate quotations. Under such challenges an investigation into virgin areas of Web Cost Estimation Techniques was undertaken. The finding of the investigation is organized into following four parts:

Part I, described in Section 2 presents an overview of Web Estimation Methods based on metrics like size, effort, cost and schedule already known.

Part II, described in Section 3 presents the Early Web Cost Estimation Model by Mendes et al.

which was the first major industrial scale investigation on the basis of huge data sets culminating into validated formula to estimate the development effort.

Part III, described in Section 4 presents the Applicability Verification Survey of the Mendes Early Web Cost Estimation Model into the Indian Context.

Part IV, described in Section 5 presents an extension of the Mendes Early Web Cost Estimation Model by considering the contingency and profitability issues and the outcome being validated have been christened into Evolving Predictor Variable Estimation Model for Web Engineering Projects.

2. SURVEY OF PREVIOUS WORKS DONE IN THE FIELD OF WEB ESTIMATON

There are two categories of applications which broadly represent the applications delivered using the Web: Web hypermedia applications and Web software applications (Christodoulou et al., 2000). A Web hypermedia application is a non-conventional application characterized by the authoring of information using nodes (chunks of information), links (relations between nodes), anchors and access structures (for navigation) and its delivery over the Web. Technologies commonly used for developing such applications are HTML, JavaScript and multimedia. These applications have great potential in areas such as software engineering (Fielding and Taylor, 2000), literature (Tosca, 1999), education (Michau et al., 2001), and training (Ranwez et al., 2000). Web software application, conversely, represents more conventional software applications that depend on the Web or use the Web's infrastructure for execution. Typical applications include legacy information systems such as databases, booking systems, knowledge bases etc. Many e-commerce applications fall into this category. Typically they employ development technologies (e.g., DCOM, ActiveX etc), database systems, and development solutions (e.g. J2EE).

2.1 Web Size Metrics for Web Cost Estimation

To date few papers have proposed Web size metrics aimed at Web cost estimation (Cowderoy, 1998; Mendes et al., 1999; Cowderoy, 2000; Mendes et al., 2000; Reifer, 2000; Rollo, 2000; Cleary, 2000; Mendes et al., 2001). Cowderoy (1998;2000), Reifer (2000) and Cleary (2000) have used industrial data sets of Web projects to justify their size metrics and to generate corresponding cost models, each collecting their data from just one Web company, possibly affecting the external validity of their results. Mendes et al. (2001) proposes size metrics for static and dynamic Web

applications and Mendes et al. (2000) proposes size metrics for Web hypermedia applications. However the data sets employed in these studies are based on Web applications developed by students, which may also affect the external validity of their results. Each of these papers is reviewed in the following sub-Sections chronologically [5][6][7][8].

2.1.1. Size Measures by Cowderoy (1998; 2000)

Cowderoy (1998; 2000) recommends several size metrics for cost estimation and risk assessment of Web application development projects. Metrics were organized by the Entities to which they apply[2].

2.1.1.1. Web application

They include *Web pages* (WP), *Home pages* (HP), *Leaf nodes* (LN), *Hidden nodes* (HN), *Depth* (DE), *Application Paragraph count* (APC), *Delivered images* (DI), *Audio files* (AF), *Application movies* (AM), *3d objects* (3DO), *Virtual worlds* (VW) and *External hyperlinks* (EH).

2.1.1.2. Web page

They include *Actions* (AC), *Page paragraph count* (PPC), *Navigational structure*s (NS), *Page movies* (PM), and *Interconnectivity* (IN).

2.1.1.3. Media

It includes *Image size* (IS), *Image composites* (ICS), *Language versions* (LV), *Duration* (DU), *Audio sequences* (AS) and *Imported images* (IMI).

2.1.1.4. Program

It includes *Lines of source code* (LOC) and *McCabe cyclomatic complexity* (MCC) (Fenton and Pfleeger, 1997).

2.1.2. Size Measures by Mendes et al.

Mendes et al. (1999; 2000; 2001) proposed size metrics to be used to predict authoring effort for hypermedia applications and then for Web applications. All metrics are presented organized by Entities to which they apply [3].

2.4.4.2.1. Hypermedia application

Hyper document size (HS): the number of documents that the hypermedia application has. Documents are considered here to be either HTML files or any kind of file that is defined as a document in the hypermedia systems used in the evaluation.

Connectivity (CON): the number of links that the hypermedia application has. These links can be internal or external. Dynamically generated links are excluded.

Compactness (Botafogo et al., 1992) (COM): measures how inter-connected the nodes are.

S*tratum* (Botafogo et al., 1992) (STR): measures to what degree the hypermedia application is organized for directed reading.

Link Generality (LG): measures if the link applies to a single instance,

for example point-to-point links, or whether it applies (or can be applied) to multiple instances.

2.4.4.2.2. Web application

Later Mendes et al. (2000) proposed a new set of size metrics, all targeting at Web applications:

Page count (PAC): the number of HTML or SHTML files used in a Web application.

Media count (MEC): the number of unique media files used in a Web application.

Program count (POC): the number of CGI scripts, JavaScript files, Java applets used in a Web application.

Total page allocation (TPA): the total space allocated for all HTML or SHTML pages (Mbytes) used in a Web application.

Total media allocation (TMA): total space allocated for all media files (Mbytes) used in a Web application.

Total code length (TCL): total number of lines of code for all programs used in a Web application.

Reused media count (RMC): the number of reused or modified media files used in a Web application.

Reused program count (RPC): the number of reused or modified programs used in a Web application.

Total reused media allocation (TRM): total space allocated for all reused media files used in a Web application (Mbytes).

Total reused code length (TRC): total number of lines of code for all programs reused by a Web application.

Code comment length (CCL): total number of comment lines in all programs in a Web application.

Reused code length (RCL): total number of reused lines of code in all programs in a Web application.

Reused comment length (ROL): total number of reused comment lines in all programs in a Web application.

Total page complexity (TPC): the average number of different types of media used in the Web application, excluding text.

Connectivity (CON): measures the total number of internal links, not including dynamically generated links.

Connectivity density (COD): computed as *Connectivity* divided by *page coun*t.

Cyclomatic complexity (Fenton and Pfleeger, 1997) (CCO): computed as *Connectivity -page coun*t) + 2.

2.4.4.2.3. The Revised List

This list was revised to include also bottom-up metrics (Mendes et al., 2001): Web page

Page allocation (PAL): measures the allocated space of a HTML or SHTML file (Kbytes).

Page complexity (PCO): the number of different types of media used on a page, not including text.

Graphic complexity (GRC): the number of graphics media used in a page.

Audio complexity (AUC): the number of audio media used in a page.

Video complexity (VIC): the number

of video media used in a page.
Animation complexity (ANC): the number of animations used in a page.
Scanned image complexity (SIC): the number of scanned images used in a page.
Page linking complexity (PLC): the number of links per page.

Media
Media duration (MED): the duration of audio, video, and animation (minutes).
Media allocation (MEA): The sizes of a media file (Kbytes).

Program
Program Code length (POL): the number of lines of code in program.

2.4.4.3 Size Metrics by Rollo (2000)
Rollo (2000) did not suggest any new size metric; however, he was the first researcher to investigate the issues of measuring the size of Web hypermedia and Web software applications, aiming at cost estimation, using several function point analysis methods. He measures the size of two applications in IFPUG, *MKII*, and *COSMIC-FFP* (Common Software Measurement International Consortium-Full Function Points) methods. Rollo (2000) concludes that *COSMIC-FFP* proved to be the most flexible approach for counting the functional size of Web hypermedia and Web software applications and can be applied to any Web application.

2.4.4.4. Size Metrics by Cleary (2000)
Cleary (2000) proposes size metrics for Web cost estimation dividing them into two types: size metrics for Web hypermedia applications and size metrics for Web software applications.

2.4.4.4.1. Web hypermedia application
Non-textual elements (NTE): the number of unique non-textual elements within an application.
Externally sourced elements (ESE): the number of externally sourced elements.
Customized infra-structure components (CIC): the number of customized infra-structure components.
Total Web points (TWP): the total size of a Web hypermedia application in Web points by adding: number of Web pages of "Low" complexity multiplied by the value for "Low" Web points; with number of Web pages of "Medium" complexity multiplied by the value for "Medium" Web points; with number of Web pages of "High" complexity multiplied by the value for "High" Web points.

2.4.4.4.2. Web software application
Function points (FPS): the functionality of a Web software application. Does not specify any particular method.

2.4.4.4.3 Web page
Non-textual elements page (NTP):

the number of non-textual elements within a Web page.

Words Page (WOP): measures the number of words in a Web page.

Web points (WPP): the length of a Web page. Scale points are "Low", "Medium" and "High". Each scale point is attributed a number of Web points, previously calibrated to a specific Web projects data set.

Number of links into a Web page (NIL): the number of incoming links; can be internal or external links.

Number of links out of a Web page (NOL): the number of outgoing links; can be internal or external links.

Web page complexity (WPC): the complexity of a Web page based upon its *number of words*, and combined *number of incoming* and *outgoing links*, plus the *number of non-textual elements*. The scale points are "Low", "Medium" and "High". Value ranges are provided for each scale point, for number of words and combination of *incoming links + outgoing links + non-textual elements*. These values have been calibrated based on a specific Web projects data set.

2.4.4.5. Size Measures by Reifer (2000)

Reifer (2000) proposes a size metric called Web Objects, which measures the number of Web Objects. Size is measured using an adaptation of Halstead's equation for volume, tuned for Web applications. The equation is as follows:

$$V = N \log_2(n) = (N1 + N2) \log2 (n1 + n2) \quad (1)$$

Where:

N = number of total occurrences of operand and operators

n = number of distinct operands and operators

$N1$ = total occurrences of operand estimator

$N2$ = total occurrences of operator estimators

$n1$ = number of unique operands estimator

$n2$ = number of unique operators estimators

V = volume of work involved represented as Web Objects

Operands are comprised of the following metrics:

Number of building blocks (NBB): number of components, e.g., Active X, DCOM, OLE.

Number of COTS (NCO): number of COTS components (including any wrapper code).

Number of multimedia files (NMM): number of multimedia files, except graphics files (text, video, sound etc).

Number of object or application points (Cowderoy et al., 1998; 2000) (NOA): the number of object or application points or others proposed (# server data tables, # client data tables etc).

Number of Lines (NLI): number of xml, sgml, html and query language lines (# lines including links to data attributes).
Number of Web components (NCM): number of applets, agents etc.
Number of graphics files (NGR): number of templates, images, pictures etc.
Number of scripts (NSC): number of scripts for visual language, audio, motion etc.

3. MENDES EARLY WEB COST ESTIMATION MODEL

All size metrics presented in the Section 2 were invariably related to implemented Web applications. Even when targeted at measuring functionality based on function point analysis, researchers only considered the final Web application, rather than requirements documentation generated using any existing Web development methods. This makes their usefulness as early effort predictors questionable. Mendes et al. (2006) conducted surveys and case study to bring light to this issue, not only by identifying early size metrics and cost drivers based on current practices of several Web companies worldwide, but also by comparing these identified metrics to those that have been proposed in the past, looking for possible convergence[4].

3.1. Mendes First Survey: Using on-line web project price quote forms

The purpose of this survey (S1) was to identify early Web size metrics and factors used by Web companies to estimate Effort for Web projects early on in the development cycle. The target population was that of Web companies that offer online Web project price quotes to customers. There was no need to contact Web companies directly, only to download their on-line Web project price quote forms from the Web. To obtain sample population a number of questions were asked by sending web forms and getting replies online.

3.1.1. Survey Results

The data collected from 133 on-line quotes was organized into six categories: Web application static metrics, Web application dynamic metrics, Cost Drivers, Web project metrics, Web company metrics, Web interface style metrics. The survey showed that out of these metrics two metrics stood out, ***total number of Web pages*** *(70% companies)* and ***features/functionality*** *(66% companies)*. Both can be taken as size metrics where the first is a typical **length size metric** and the second an abstract measure of **functionality**. Seventy four (74) Web companies were also asked for the available Web project budget. Mendes et al. believe this metric can have a bearing on the contingency and/or profit costs that are provided

in a price quote. Project estimated end date, project estimated start date and application type also were important. Mendes et al. believe these help set priorities and perhaps decide on what skills are necessary and available to the project.

3.1.2 Case Study for validating the results obtained from Survey

The survey identified metrics related to a project's price quote. Mendes et al. applied to their work the same model employed in Kitchenham et al., 2003, where price is the results of three components: estimated effort, contingency and profit. Since their objective was to identify only those metrics specifically targeted at effort estimation, they employed a case study and a second survey to identify the subset of metrics obtained in first survey directly related to effort estimation. The case study consisted of contacting an experienced Web company to confirm/deny, based on the ranking provided, those metrics they consider important for early Web cost estimation. The Web Company contacted is based in Rio de Janeiro, Brazil and five people work in Web design and development within the company and has to date a portfolio of more than 50 Web applications. They document their development process and record size and effort metrics from past projects. Depending on the type of project, one out of two types of process models is used: prototyping or waterfall. The choice depends on the Company's familiarity with the application domain. The Company's effort estimation practices are based solely on expert opinion, where the average estimation accuracy for their Web projects, based on effort estimates obtained early in the life cycle, is 10%.

The company's director was asked to help validate those metrics obtained from first Survey and validation here represents identifying size metrics and cost factors important to be used in the Web cost estimation process early in the development life cycle. Therefore all metrics from earlier survey selected by the Company director were actually employed on their Web cost estimation process. This person has worked in software development for more than 20 years and is experienced in management of large projects, conventional or Web-based. For Web application static metrics agreement was reached for most metrics. For Web application dynamic metrics more features/functionality were added to the list and the director suggested that adding a complexity level to each feature/functionality would help discern more difficult implementations. More specifically, this Company groups functions/features within three groups: simple, complex and very complex. Each has an associated

baseline, which represents a percentage to be added to estimated effort. The Company's baselines reflect average percentages based on past experience. These metrics were also confirmed as suitable for early Web cost estimation. The metrics Mendes et al. have obtained as a result of their first survey and validated by the mature Web Company corroborate their findings.

3.2 **Second Survey for validating the results obtained from First Survey**

The purpose of this second survey was also to validate the results that have been obtained from first survey. Mendes et al. have also considered in this survey some of the results obtained from the case study, more specifically regarding Web application dynamic metrics. The target population was that of Web companies in New Zealand that estimate effort for their Web projects. The survey instrument was a questionnaire with nine questions was prepared by one of the team members, and the method of gathering data was via interviews over the phone. The results they obtained validated to a large extent the results obtained from the first survey.

3.3. **Tukutuku Benchmarking Project**

The feedback obtained from the first and second Surveys and the case study was used by Mendes et al. to prepare Web forms to gather data on Web projects worldwide. This data gathering initiative was called the Tukutuku benchmarking project11. The Tukutuku project aimed at gathering data on Web projects worldwide to be used to develop Web cost estimation models based on early effort predictors and to benchmark productivity across and within Web Companies. While preparing the project data entry forms used in Tukutuku, careful consideration was given to differentiate more complex features/functions from less complex ones, as this was then current practice of some of the Web companies surveyed in the second survey and also by the mature Web Company from their case study. They had a detailed list of features/function obtained from the first survey and the case study. Although some certainly seemed more complex than others they did not want to suggest or impose any complexity, leaving it for each Web company to decide. The solution was devised as follows:

- Companies contributing Web project data to Tukutuku were asked to indicate (tick) all the features/functions that the application had.
- For each feature/function they would also indicate if it was a black box reuse, reuse with adaptation, or new development.

- They were also asked to indicate if a given feature/function employed high effort to be developed / adapted.
- To be familiar with what each Company understood by high effort they also asked them to indicate the effort in person hours that would be automatically representative of high effort to develop or adapt a feature/function.
- Companies would be able to provide features/functionality that they had not considered.

The Tukutuku Benchmarking project started in 2006 and till 2006 had received 67 Web projects from 25 Web companies in 9 different countries. 27 projects came from two companies (13 and 14 respectively). Each Web project in the database provided 43 variables to characterize a Web application and its development process. Mendes et al. were aware that the data obtained was a result of a self-selected sample. However they believed the data in the Tukutuku database can be very useful as an indicator provided one is aware of the limitations. No automated measurement tools were used by the Web companies that volunteered data for the Tukutuku database. Therefore the accuracy of their data could not be determined. In order to identify guesstimates from more accurate effort data, they asked companies how their effort data was collected. They found that in at least 77.6% of Web projects in the Tukutuku database effort values were based on more than guesstimates. However, Mendes et al. are also aware that the use of timesheets does not guarantee 100% accuracy in the effort values recorded. The data collected to date for the Tukutuku project has not followed rigorous quality assurance procedures to validate the data and the projects' application domains are mixed.

3.4. **Using Multivariate Regression to Identify Early Web Size Metrics and Cost Factors**

All the variable data were analyzed using multivariate forward stepwise regression. The set of variables used for building the cost models is shown in Table 1. This is a subset of the Tukukuku data set since several variables had to be excluded if they were within the constraints of most instances of a variable being zero, the variable was categorical or the variable was related to another variable, in which case both could not be included in the same model. This was investigated using a Spearman's rank correlation analysis ($\alpha = 0.05$). Whenever variables were highly skewed they were transformed to a natural logarithmic scale to approximate a normal distribution (Maxwell, 2002). In addition, whenever a variable needed to be transformed but had zero values, the natural

logarithmic transformation was applied to the variable's value after adding 1.

TABLE1: VARIABLES USED IN STEPWISE REGRESSION

Variable	Meaning
lntoteff	Natural log. Of the total effort to develop a Web application.
nlang	Number of different languages used on the project
devteam	The number of people who worked on the project
teamexp	Average team experience with the development language(s) employed
lnnewwp	Natural log. of (1+ number of new Web pages)
lnimgnew	Natural log. of (1+ number of new images in the applications)
lnimglib	Natural log. of (1+ total number of images reused from a library)
lnimg3p	Natural log. of (1+ total number of images developed by a third party)
hfotsa	Total number of adapted high effort functions.
lntoth	Natural log. of (1+ total number of high effort functions).
totnhigh	total number of low effort functions
Natural log. = Natural logarithm	

The final Mendes Early Web Cost Model is presented in Table 2 below.

TABLE 2 : BEST FITTING MODEL TO CALCULATE lntoteffor

Independent Variables	Coefficient	Std. Error	t	p>|t|	95% Confidence Interval
(constant)	2.154	0.260	8.281	0.000	1.634 – 2.674
lnnewwp	0.435	0.061	7.184	0.000	0.314 – 0.556
lntoth	0.671	0.160	4.198	0.000	0.352 – 0.991
devteam	0.239	0.083	2.876	0.005	0.073 – 0.406

The equation as read from the final model's output is:

ln(toteffor) = 2.154 + 0.435 × ln(newWP+1) + 0.671 × ln(tothigh+1) + 0.239 × devteam . (1)

Which, when transformed back to the raw data scale, gives the equation:

toteffor = 8.619 × (newWP+1)0.435×(tothigh+1) x 0**.671 × *e*0.239 × devteam (2)**

Where, toteffor is total effort, newWP is new web page, tothigh is total high functions and devteam is size of the development team.

When there is no reuse, Webpages is the same as newWP.

Despite this cost model not presenting good estimation accuracy, its main objective is to indicate that two of the variables selected by the best fit model are a very close match to the two variables ranked highest in the first survey – number of Web pages and number of high effort features/functions. Mendes et al. believe this result to be very promising, suggesting that these metrics can be estimated by customers early in the development life cycle and are suitable for building Web cost models at, for example, the bidding stage. It is to be noted that to obtain the number of high effort functions it is necessary to provide the client with a list of high effort functions and to count those that have been selected. In addition, the development team size must be already known by the Web Company.

4. APPLIABILITY VERIFIATION SURVEY OF THE MEDES EARLY WEB COST ESTIMATION MODEL INTO THE INDIAN CONTEXT

The Indian Software industry started flourishing in the last decade of twentieth century thanks to easy availability of software personnel and outsourcing of software development work from the United States. In the first decade of the present century it emerged as a major player with entry of Multinational Corporations into Indian soil and diversification of the domestic Indian Software industrial houses into multinational operations. Today India is recognized as a Software Giant all around the world and its software exports has expanded beyond imagination. We now produce software relating to business applications and insurance, banking and financial applications, multimedia, avionics and space research, educational and instructional, e-governance and service delivery, enterprise information management and ERP solution, e-commerce and m-commerce, mobile computing and cloud computing, database management and system development etc. and a vast majority of them are web application

development. They cater not only to the domestic market but also to international trade requirements. No doubt some of the Indian software companies that started from scratch about 25 years ago have now turnover exceeding billions of $. It was therefore necessary to streamline the Indian software development industry and this could be done by providing some help to them by evolving a cost estimation model based on more realistic variables than Mendes prescribed and more valid in the Indian context.

4.1. Limitations of Mendes Early Web Cost Estimation Model

Mendes et al. have only hinted at the contingency and profitability as cost variable but neglected them in their calculations. This has made their work a little bit approximate than would have been if both of them were also taken into account. It was felt necessary to investigate further into the matter to ameliorate the results obtained by them.

4.2. Verification Survey of Indian Software Projects

A survey was felt necessary to be conducted with Software projects completed by Indian Companies. In the beginning India's Top Twenty Software Companies were contacted in the year 2011 by means of formal request letters over speed post. Only two companies replied to cooperate in this survey. It was then decided to follow up with telephone and mobile over which they replied either to contact their software vendors or that they are simply not interested and started questioning the very intent of this survey. When contacted with their vendors two more vendors agreed to provide the data sets of their completed projects. Two more small local companies were also persuaded to part with such data sets such that the company size becomes homogenous. The following data were obtained with much persuasion:

TABLE3: DATA GATHERED DURING THE APPLICABILITY VERIFICATION SURVEY OF INDIAN PROJECTS.

S .N.	Project type	New WP (No.)	Tothigh (No.)	Dev Team (No.)	Toteffor (hrs-worked)	Crim (hrs-worked)	Grosseffor (hrs-worked)
1	Business app.	5	11	3	215	60	275
2	Business app.	18	61	6	2138	438	2576
3	Busin.ess app	12	16	8	1192	132	1324
4	Educational	17	22	9	2308	378	2686
5	multimedia.	28	48	10	5792	1043	6835
6	Financial	17	54	11	6308	625	6933
7	e-governance	22	59	12	9420	1188	10608
8	Banking	34	89	14	22920	3811	26731
9	Share market	47	126	15	44570	5432	50002

Where newWP, Tothigh, devteam, toteffor have usual meanings as per equation (2) and crim is contingency and risk management effort and grosseffor is gross effort of development including crim.In the table3 above projects have been arranged in the increasing order of development team size. Further data were gathered relating to development time; cost of development and profitability etc .which has been arranged in table 4 as follows:

TABLE 4: COST AND PROFITABILITY DATA

S.N.	Project Id	Hrs. per day	No. of days	Dev. Cost in INR (C1)	Crim cost in INR (C2)	Cost price CP in INR (C1+C2)	Sale price SP, in INR	Profit in INR	Profit in %
1	P1	10	9	17080	4805	21885	25000	3115	14.2
2	P2	10	43	116667	21020	137687	163699	26012	18.8
3	P3	10	17	47212	5115	52327	61540	9213	17.6
4	P4	14	21	110210	18851	129061	143379	14318	11.1
5	P5	8	83	214502	31892	246394	273710	27316	11.1
6	P6	12	58	382120	78210	460330	505000	44670	9.7
7	P7	11	80	603500	76500	680000	742125	62125	9.1
8	P8	12	140	1232808	283360	1516168	1736013	219845	14.5
9	P9	12	248	2645330	320000	2965330	3298900	333570	11.2

4.3.1. Verification

Calculation

The above data were analyzed further in the light of Mendes Early Web Cost Model as per the following:

4.3.1.1. Project P1

As per Mendes Eqn.(2),

$$\text{toteffor} = 8.619 \times (\text{newWP}+1)^{0.435} \times (\text{tothigh}+1)^{0.671} \times e^{0.239 \times \text{devteam}}$$

$$=8.619*(5+1)^{0.435}*(11+1)^{0.671}*e^{0.239*3}$$

$$=8.619*2.181*5.298*2.048$$

=203.96 hours worked.

As per observed data,

Grosseffor= 9 days of 3 persons working @10 hrs/day

=9*3*10 hrs

= 270 hrs.

Mean Payment/ hr. = CP/grosseffor

=21885/275

=89.58

Development time =C1/mean payment rate

=17080/89.58

=190.67 hrs.

For 3 persons@10 hrs./day, dev. time in No. of

days

=190.67/3*10

=6.35 days

Additional time for contingency and risk management, crim,

=crim cost/mean payment rate

=4805/89.58

=53.64 hrs.

For 3 persons@10 hrs./day, dev. time in No. of days

=53.64/3*10

=1.79 days

Total deveop. time =6.35+1.79

=8.14 days

4.3.1.2. Project P4

As per Mendes Eqn.(2),

$\text{toteffor} = 8.619 \times (\text{newWP}+1)^{0.435} \times (\text{tothigh}+1)^{0.671} \times e^{0.239 \times \text{devteam}}$

$=8.619*(17+1)^{0.435}*(22+1)^{0.671}*e^{0.239*9}$

$=8.619*3.516*8.198*8.593$

=2134.8 hours worked.

As per observed data,

Grosseffor= 21 days of 9persons working @14 hrs/day

=21*9*14 hrs

= 2646 hrs.

Mean Payment/ hr. = CP/grosseffor

=129061/2646

=48.78

Development time =C1/mean payment rate

=110210/48.78

=2259.33 hrs.

For 9 persons@14 hrs./day, dev. time in No. of days

=2259.33/9*14

=17.93 days

Additional time for contingency and risk management, crim,

=crim cost/mean payment rate

=18851/48.78

=386.45 hrs.

For 9 persons@114 hrs./day, dev. time in No. of days

=386.45/9*14

=3.07 days

Total deveop. time =17.93+3.07

=21 days

4.3.1.3. Project P9

As per Mendes Eqn.(2),

$$\text{toteffor} = 8.619 \times (\text{newWP}+1)^{0.435} \times (\text{tothigh}+1)^{0.671} \times e^{0.239 \times \text{devteam}}$$

$=8.619*(47+1)^{0.435}*(126+1)^{0.671}*e^{0.239*15}$

=8.619*5.387*25.802*36.053

=43191.53 hours worked.

As per observed data,

Grosseffor= 248 days of 15 persons working @12 hrs/day

=248*15*12 hrs

= 44640 hrs.

Mean Payment/ hr. = CP/grosseffor

=2965330/44640

=66.43

Development time =C1/mean payment rate

=2645330/66.43

=39821.31 hrs.

For 15persons@12 hrs./day, dev. time in No. of days

=39821.31/15*12

=221.23 days

Additional time for contingency and risk management, crim,

=crim cost/mean payment rate

=320000/66.43

=4817.1 hrs.

For 15 persons@12 hrs./day, dev. time in No. of days

=4817.1/15*12

=26.76 days

Total develop. time =221.23+26.76

=247.9

9 days

TABLE 5: PREDICTION ACCURACY OF DATA ANALYZED

S.N.	Project Id	Prediction			Observation			Accuracy %
		Toteffor in hrs. worked	Crim in hrs. worked	Grosseffor In hrs. worked	Toteffor in hrs. worked	Crim in hrs. worked	Grosseffor In hrs. worked	
1	P1	203.96	53.64	257.6	215	60	275	6.32
2	P4	2134.80	386.45	2521.25	2306	378	2686	6.13
3	P9	43191.53	4817.1	48008.63	44570	5432	50002	3.99

In the above analysis the initial assumption that contingency and risk management is a cost variable has been proved to be valid as the errors are well within the permissible limits and are arising because of various assumptions made so as to simplify web cost effort estimation.

5. Evolving Predictor Variable Estimation Model for Web Engineering Projects

With a view to accommodate the contingency and risk management cost into the Mendes Equation it was decided to investigate further into the data observed and predicted and evolve a predictor variable estimation model of web engineering projects by using statistics and calculus.

5.1 Using Multivariate Regression to derive Predictor Variable Web Estimation Model

5.1.1 Curve Fitting by Sum of Exponentials

In the analysis of web development data we can use the fitting of a sum of exponentials of the form,

$$y = A_1e^{\lambda 1x} + A_2e^{\lambda 2x} + A_3e^{\lambda 3x} + \ldots\ldots\ldots\ldots + A_ne^{\lambda n} \quad (3)$$

Where $A_1, A_2, A_3 \ldots\ldots A_n$ and $\lambda_1, \lambda_2, \lambda_3, \ldots. \lambda_n$ are unknowns.

Eqn. (3) satisfies a differential equation of the type:

$$d^ny/dx^n + a_1 d^{n-1}y/dx^{n-1} + a_2 d^{n-2}y/dx^{n-2} + a_ny = 0 \quad (4)$$

Where $a_1, a_2, a_3 \ldots a_n$ are unknowns.

5.1.1.1. Froberg's Method

Froberg suggested a method for computing these derivatives numerically in equation (4) at the given points and substituting them in equation (3), thus obtaining a system of n linear equations for n unknowns $a_1, a_2, a_3 \ldots a_n$ that can be solved.

Again it can be seen that $\lambda_1, \lambda_2, \lambda_3 \ldots \lambda_n$ are the roots of the polynomial equation.

$$\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots\ldots\ldots + a_n = 0 \quad . (5)$$

Which when solved enables us to determine A_1, A_2, A_3, …. A_n, from equation (3) by method of least squares. An obvious disadvantage is with their increasing order and therefore leading to unreliable results.

5.1.1.2 Moore's Method

Equation (4) can be solved by Moore's method which is described below for n =2.

For n = 2, eqn. (3) becomes,

$$y = A_1e^{\lambda 1x} + A_2e^{\lambda 2x} \quad (6)$$

This satisfies the differential equation,

$$d^2y/dx^2 = a_1dy/dx + a_2y \quad (7)$$

Assuming a is the initial value of x and integrate eqn. (7) w.r.t. x we get,

$$y'(x) - y'(a) = a_1 y(x) - a_1 y(a) + a_2 \int_a^x y(x)dx \quad . \quad (8)$$

Where, $y'(x) = dy(x)/dx$.
Integrating equation (8) again w.r.t. x from a to x, we will get,

$$y(x) - y(a) - (x - a)y'(a) = a_1 \int y(x)d(x) - a_1(x-a)y(a) + a_2 \int \int y(x)d(x)d(x) \quad . \quad . \quad (9)$$

Using results from calculus, we obtain,

$$\int \ldots \int f(x)dx \ldots dx = \frac{1}{|\_(n-1).} \int (x-t)^{n-1} f(t)d(t) \quad . \quad . \quad . \quad . \quad (10)$$

Equation (10) simplifies to,

$$y(x)-y(a)-(x-a)y'(a) = a_1 \int y(x)dx - a_1(x-a)y(a) + a_2 \int (x-t)y(t)dt \quad . \quad (11)$$

In order to use eqn. (11) to set up a system of linear equation in terms of a_1 and a_2, $y'(a)$ need to be eliminated and this is done in following way:
Choosing two data points x_1 and x_2 such that $a-x_1=x_2-a$, then from eqn.(11) we will get,

$$y(x_1)-y(a)-(x_2-a)y'(a) = a_1 \int y(x)dx - a_1(x_1-a)y(a) + a_2 \int (x_1-t)y(t)dt \quad (12)$$

$$y(x_2)-y(a)-(x_2-a)y'(a) = a_1 \int y(x)dx - a_1(x_2-a)y(a) + a_2 \int (x_2-t)y(t)dt \quad (13)$$

Again simplifying the eqns. (12) and (13) by using $a-x_1=x_2-a$, we get,

$$y(x_1)+y(x_2)-2y(a)=a_1[\int y(x)dx + \int y(x)dx] +a_2[\int (x_1-t)y(t)dt+\int (x_2-t)y(t)dt] \quad . \quad . \quad (14)$$

Now eqn.(14) can be used to setup a system of linear equations for a_1 and a_2 and then we obtain λ_1 and λ_2 from characteristic equation.

$$\lambda^2=a_1 \lambda+a_2 \quad . \quad . \quad . \quad (15)$$

Finally, A_1 and A_2 can be obtained by the method of least squares. Thus we obtain the required form of the equation as,

$$y=A_1 e^{\lambda 1 x} + A_2 e^{\lambda 2 x} \quad . \quad . \quad (16)$$

5.2 Results Obtained by Using Curve fitting by Sums of Exponentials

In order to use curve fitting by sums of exponentials let us neglect the profitability issue and concentrate on total development effort and

contingency and risk management effort. Thus the order of the differential equation reduces to 2 and then Moore,s Method can be applied on the data sets. By using curve fitting and solving differential equation on the observed data set we get the following equation of gross effort and current sale price of the web projects,

$$grosseffort = 102.74\,(e^{0.41*devteam} - e^{-3.34*devteam}) \quad . \quad . \quad . \quad (17)$$

and,

$$C.S.P. = 10043*\mathbf{(e^{0.41*devteam} - e^{-3.34*devteam})} \quad . \quad . \quad (18)$$

5.3 Validating Evolving Predictor Variable Web Estimation Model

To validate the estimation model derived using the predictor variable equation (17) and (18) for independent project P10 with following observed parameters:

Test Case 1

Project Id = P_{10}

Development team in no., (devteam) = 8,

Gross effort = 2548 hrs. worked.

CSP in INR = 246500 at the current(2014) price level.

Comparing the validation set with the results obtained from the proposed model:

Project Id = P_{10}

Development team in no., (devteam) = 8,

Predicted Gross effort = $102.74*(e^{0.41x8} - e^{-3.34x8}) = 2732.9$ hrs. worked.

CSP in INR = 266900 at the current (2014) price level.

Gross effort = $102.749*(e^{0,41*3}\text{-}6185e\text{-}3.34^{*3})$

$= 5696*1.481\text{-}6185*1.318$

$= 8435.78\text{-}8151.83$

$= 283.95$ hrs. worked.

Accuracy = 275-283.95/275

= 3.25% hence well within tolerable limits.

On an average it was found that,

The gross effort prediction accuracy is 1.7%.

The current sale price prediction accuracy is 7.6%.

As shown in graph below:

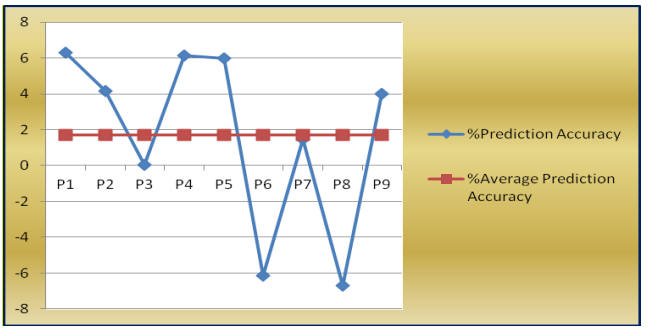


Fig.1: Line Chart Showing the variation of prediction accuracy in % of various projects as compared to the Average prediction accuracy.

6. SUMMARIES AND CONCLUSION

The paper highlighted the prevalent software estimation models and Mendes Early Web Cost Estimation Model in detail and the need to extend it by considering the additional issues of contingency, risk management and profitability issues so as to be more accurate and realistic. Thus by using tools like regression and calculus and it derived the Evolving Predictor Variable Estimation Model for Web Engineering Projects on the basis of the following equations:

grosseffort = 102.74 ($e^{0.41*devteam} - e^{-3.34*devteam}$) in hours worked.

C.S.P. = 10043*($e^{0.41*devteam} - e^{-3.34*devteam}$) in INR(at 2014 price).

Where devteam is the number of team members.

It would be very useful for web application development projects in early gross effort and cost prediction [9].

7. ACKNOWLEDGEMENT

The authors wish to acknowledge all the companies, software professionals, friends and faculties who generously contributed to the research work failing which this project would not have been possible.

REFERENCES

Srivastava, Sanjay Kumar & Varma, S.P. (2012). Software Projects Estimation Models, *IT Floor, Vol.1, No.2, pp.134-140,* B.R.A. Bihar University, Muzaffarpur, Bihar, PIN-842001

Cowderoy, A.J.C. (2000), Measures of Size and Complexity for web site content, *Proc. Combined 11th ESCOM conference and the 3rd SCOPE Conference on Software Product Quality, Munich, Germany, pp.423-431.*

Mendes, E., Mosley N. & Counsell, S. (2002) Comparison of Length, Complexity and functionality as Size Measures for Predicting Web Design and Authoring Effort, *IEEE Proc. Software pp. 149.3.86…92.*

Mendes, E., Mosley, N. & Counsell, S. (2007) Investigating Early Web Size Measures for Web Cost Estimation, *IEEE Proc. Software.*

Mendes, E. & Kitchenham, B.A. (2004) Further Comparison of Cross-company and within-company Effort estimation Models for Web Applications, *submitted manuscripts.*

Mendes, E. & Mosely, N. (2000) Web Metrics and Development Effort Prediction, *Proc. ACOSM, 2000, Sydney Australia.*

Mendes, E., Hall, W. & Harrison, R. (1999) Applying measurements principles to improve hypermedia authoring, *New Review of Hypermedia and Multimedia, Taylor Graham Publishers, Chapter 5, pp.105-132.*

Mendes, E., Mosely, N. & Counsell, S. (2001) Web Metrics – Estimating Design and Authoring Effort, *IEEE Multimedia, Special Issue on Web Engineering, Jan.-Mar., 2001, pp.50-57.*

Srivastava, Sanjay Kumar & Varma, S.P. (2014) Evolving Predictor Variable Estimation Model for Web Engineering Projects, *Proc. of the 1st international Conference on Recent Trends in Computer Science and Engineering, Central University of Bihar, Patna, 08-09 Feb. 2014, pp.335-346.*