Abdoulaye Gamatié *Editor*

Computing in Research and Development in Africa

Benefits, Trends, Challenges and Solutions

Springer

Computing in Research and Development in Africa

Abdoulaye Gamatié
Editor

Computing in Research and Development in Africa

Benefits, Trends, Challenges and Solutions

Springer

Editor
Abdoulaye Gamatié
Laboratory of Informatics,
Robotics and Microelectronics
of Montpellier
Montpellier, France

ISBN 978-3-319-08238-7 ISBN 978-3-319-08239-4 (eBook)
DOI 10.1007/978-3-319-08239-4
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014948198

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The ability of Africa to sustain a stronger development in the twenty-first century certainly depends in part on its adequate adoption of information and communication technology (ICT) for defining innovative solutions to crucial challenges it recurrently faces with regard to disease, hunger, education, unemployment, and transportation. While some existing initiatives already exist in that direction, there are still many opportunities for Africa in exploiting the power of computing and mobile communication to deal with such challenges so as to reach a competitive level comparable to developed countries. The rapid spread of ICT currently observed in all parts of Africa is one major reason to be optimistic in its high capacity to embrace modern technology. This is in my opinion an unprecedented opportunity to set up a continent-level synergy for exploring common solutions to the crucial problems.

The main motivation of this book is to contribute to this synergy by:

- promoting a continent-level vision on the use of computing and its benefits to deal with crucial problems in developing countries, and particularly in Africa. This is a unique attempt to provide such a global vision;
- identifying the most important trends, challenges, and solutions under consideration by very active and relevant scientific researchers on these topics;
- stimulating a better structuration of research and interaction between scientific researchers on the problems faced by developing countries, and in Africa particularly;
- presenting the current gained insights related to all addressed relevant topics so that both governments and international funding organizations can more efficiently orientate their strategic decisions towards the most important development sectors.

While stressing the benefits of computing for several important domains in Africa, the book presents many important initiatives currently conducted on the continent regarding scientific research and development. In particular, a thorough

and original coverage is given on the trends, challenges, and solutions regarding computing usage within a number of areas. The book is mainly composed of four parts as follows: (1) natural resources and health management, (2) social networks and economy, (3) innovative technologies with focus on embedded systems, and (4) new governance models.

The contributed chapters present continent-level representative works proposing solutions to the identified issues by considering computing as key ingredient. They come from recognized and active field experts in research and its application with regard to the aforementioned domains in Africa. This makes the book well suited for learning and understanding the numerous development stakes and attractive opportunities in developing countries more generally speaking, beyond Africa.

More precisely, the book targets a wide audience comprising:

- Academic actors: it provides an excellent state of the art on the usage of computing to address typical problems in developing countries, pertaining most important domains. It will certainly help interested researchers as well as graduate students to better understand the current hot issues, and how the scientific research is currently structured around the African continent in order to cope with the issues. This is very important to know for setting up adequate inter-African (and beyond) research collaborations for more efficient results.
- Industry actors: the contributing authors bring a number of insights and interesting suggestions that will be useful to future investors in Africa regarding the use of modern numerical technologies in addressed domains. Indeed, with the notable progress margin still observed in the establishment of such technologies in Africa, the book draws many indications on promising sectors for investment.
- Governments, international funding organizations, and other decision makers: it is now unquestionable that information technology is essential to the development of developing countries. Government officials, international funding organizations and other decision makers will find here very relevant information for a better steering of their strategic decisions. This will result in a more efficient impact on populations, environment, and economy on the African continent, and in other developing countries beyond Africa.

My closing words concern acknowledgements. First, I would like to warmly thank all authors who kindly accepted to contribute[1] to this book upon my invitation. Via these contributions, the readers of this book will get a unique up-to-date and representative coverage of addressed topics from most parts of Africa, especially from the following countries: Algeria, Cameroon, Egypt, Ghana, Kenya, Namibia, Nigeria, Senegal, South-Africa, Tunisia, and Uganda. Many thanks to Vianney

[1]Unfortunately, it was not possible to include a few expected chapters on education, renewable energy, and high-performance computing due to very tight delivery deadlines for their authors.

Lapôtre and Rabie Ben Atitallah who participated in the review of the presented contributions. Finally, I am grateful to Charles Glaser and Jessica Lauffer, both from Springer, for their support in the realization of this book.

Montpellier, France Abdoulaye Gamatié

Contents

Part I
On Natural Resources and Health Management

Framework for Time Relevant Water Monitoring System

F.A. Katsriku, M. Wilson, G. Gyamfuah Yamoah, J-D Abdulai, B.M.A. Rahman, and K.T.V. Grattan

Abstract The emergence of telecommunications in the last decade and advances in the field of instrumentation have enabled many new applications to be developed in particular for remote monitoring of physical environment by sensors spatially distributed. A framework for water quality monitoring is been proposed based on a telecommunications infrastructure. In this proposal, we review some of the benefits that may be derived by developing nations and key challenges that need to be addressed in the monitoring of water quality for rural communities in Africa. It is proposed that Fibre Optic Sensors can be used to design compact and highly efficient sensor systems. The proposed framework may be extended to other areas of national economic importance.

1 Introduction

The developing world, Africa especially, has many challenges, infrastructural provisioning being one of the most critical. In most developing countries, the infrastructure required to deliver some of the basic services of any modern day society is inadequate. Water which is essential to the sustainability of all forms of life ranks as one of the most basic human requirements. According to the World Health Organisation, five out of the six people, without access to an improved water source live in rural areas [19]. Community water supplies in developing countries are more frequently associated with outbreaks of waterborne disease than urban supplies. According to the World Health Organization, over one billion people do not use water from an improved source [19]. The World Health Organization estimates that there are over two million deaths per year directly attributable

F.A. Katsriku (✉) • M. Wilson • J-D Abdulai
Computer Science Department, University of Ghana, Accra, Ghana
e-mail: fak30@cam.ac.uk; fkatsriku@ug.edu.gh

G.G. Yamoah
Department of Engineering, All Nations University College, Koforidua, Ghana

B.M.A. Rahman • K.T.V. Grattan
School of Engineering and Mathematical Sciences, City University, London, UK

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_1

to unsafe water. Over 20 diseases are related to water and sanitation amongst which the following can be listed; diarrhoea—over two million deaths annually, schistosomiasis—an estimated 260 million affected worldwide. Cholera is still reported in over 50 countries [18]. World Health Organization again estimates that 4 % of the global health burden could be prevented by improving water supply and hygiene [18]. A significant proportion (75 %) of the drinking water used in Africa comes from groundwater and this is often used without any treatment. Detailed knowledge of water quality is required to enable proper treatment and prevention of contamination. In Africa, women and children walk long distances daily to fetch water, mainly from an unsafe source. When this water source becomes contaminated, authorities are unaware and it is only when deaths begin to occur that the local population might suspect the water source as the problem. Due to the poorly developed infrastructure that currently exists, response to outbreaks of waterborne disease can be very slow. The response is typically to dispatch a team to the area. The team returns with water samples which are then analysed in the laboratory to identify the problem, only then would the necessary action be taken. All this is after deaths have already occurred. Even then, the results might not be accurate given the temporal variation in the water condition. If the authorities can be provided with timely information on an impending health-related situation they would be able to intervene to prevent an outbreak of a disease and the problems that follow from that.

Investing in these water supplies is crucial to improving the well-being of the population of the developing world: it will reduce waterborne disease outbreaks and the overall costs and suffering associated with that. Progress in providing safe source of drinking water for these communities has been slow to date but the proposed methodology aims to take a significant step forward by creating a framework water quality monitoring systems.

Whilst there is no physical shortage of water on earth, the majority of it, 97 %, is saline. Access to safe drinking water varies significantly and depends on the geographic location. One of the greatest challenges facing mankind is how to manage, in a sustainable way, our water resources. A key element in the management of water resources is quality monitoring. There are significant technological challenges to enable accurate real-time monitoring of water quality across a nation or even across a continent.

To understand the quality of water it is essential that it is monitored. Monitoring is defined by the International Organization for Standardization (ISO) as: “the programmed process of sampling, measurement and subsequent recording or signalling, or both, of various water characteristics, often with the aim of assessing conformity to specified objectives” [7]. Particularly in Africa, where the provisioning of safe drinking water is inadequate, it is even more so important. The ability to monitor the quality of water with a high temporal granularity will enhance greater understanding of the biological and chemical composition and enable effective remedial action to be taken. One of millennium development goals (MDG 7) is to halve by 2015 the proportion of people without sustainable access to safe drinking water [16]. Indications are that while the rest of the world is

on target to achieving this goal, progress in sub-Saharan Africa is slow. In many African countries, responsibility for water resource management is spread across many departments and thus becomes an obstacle to improving access to improved water source.

2 Background

The detection of pathogenic micro-organisms in water to prevent infection, illness and economic loss has always been a challenge. In 2000, economic research service estimated that the cost associated with five major pathogen including *Salmonella* and *Escherichia coli* amounted to at least US$7 billion annually [12]. Therefore there is a serious and pressing need for novel, rapid and more practical techniques for identification and/or monitoring of bacterial contamination in water. Photonic and nanophotonic sensors are emerging as very attractive devices to be employed in a great number of application fields such as environment, microbiology, medicine and many more [13, 20]. These optical sensors are of major interest because they exhibit shorter response time and higher sensitivities.

It is generally accepted that the traditional method of monitoring water quality through spot sampling and laboratory analysis is highly unlikely to provide a true reflection of the concentration of particular physicochemical or microbial variables present in the water body due to the very high temporal and spatial variability that characterize water bodies. With the rapidly developing telecommunications infrastructure in Africa, it has become possible to integrate telecommunications with novel methods in instrumentation based on photonics and optical techniques to deliver timely information to enable a range of actions that enhance the well-being of the population to be taken. A framework for time relevant water quality monitoring system is being proposed. The proposed framework even though has a focus on water quality could be adapted for other parameters of interest. The specific aims of this framework to enhance the quality of potable water by deploying affordable sensors in selected areas will test the quality of drinking water; the information collected by these sensors will be sent by wireless link through the existing telecommunication infrastructure to a server. The data collected can then be processed and the information made accessible through an internet service, thereby creating a central monitoring system. Key parameters to be tested may include pH level, temperature, total coliform, conductivity, colour, total dissolved solids (TDS), total suspended solids (TSS), turbidity, dissolved oxygen and chemical contamination at the chosen locations. The outcome of this work will be a system designed to enable the authorities to be alerted automatically to any contamination in drinking water source and interventionist action taken. The system, when developed, could be readily extended to deal with other areas of the national economy such as health, agriculture, environmental protection and industry. The proposed framework can be regarded as an essential element in the realization of a nationwide framework

for central monitoring, providing real-time and continuous monitoring of spatio-temporal distribution of key water quality parameters or other parameters of interest.

Traditional methods of testing for water quality involve sample and grab technique. As indicated earlier, this method is quite expensive and may not provide an accurate reflection of the current situation that pertains due to the temporal variation which is inherent in the water quality parameters. As such, this is not often undertaken except where a problem has been reported. To overcome some of the problems associated with traditional methods it has been proposed to monitor water quality in situ using portable automated instrumentation. This however does not overcome the problem of costs associated with each field visit leading to limited data points.

The current proposal is based on the development of wireless sensor networks (WSNs) to monitor, in real-time, water quality parameters. The proposed system will generate large amounts of data that scientists will find very useful for their research work. Current monitoring systems do not generate enough data to be very useful for research and other purposes. The United Nations Global Environmental Monitoring System (GEMS)/Water system is a clear demonstration of the limitations of current systems of monitoring [17]. In Ghana as an example, four monitoring stations were established in 1991. The last time data was sampled from them in 1995. The total number of data points available from the four stations is 2,127. In Uganda, 17 monitoring stations were established in 1978 of which only seven were operational. The last time data was collected from these stations in 1980. The total number of data points available from these stations is 2,858. The use of sensor networks with a large number of nodes with automatic data collection capabilities will increase the resolution of scientific data by orders of magnitude. The framework being proposed will help African scientists to begin to evaluate new and developing technologies, create training programmes, develop collaborations and participate in joint efforts with experts in sensor technology, communications, information management, and networking to design and implement prototype sensor networks. The use of WSN can also enhance educational curricula in academic institutions as well as provide local scientists with high quality data that can be analysed. The design and deployment of a WSN requires a team-based, integrative approach. Long-term data from sensor networks will be valuable for educational use, and tools and development of curricula should be encouraged.

In Ghana like in other parts of Africa, groundwater forms a significant source of domestic water source. In addition to biological contamination, other problems associated with groundwater in Ghana are excess Iron; often significant in all aquifers, excess manganese in many locations across the country. Fluoride is also found in excess quantity (>4 mg/L) in the upper regions. Arsenic is also found in excess quantity (>0.01 mg/L) mainly in the south west. The Water Research Institute (WRI) is one of the bodies responsible for water quality monitoring in the country. With aid from the Danish Government, monitoring stations have been established in Accra and Tamale. The objective of these laboratories is to perform analysis of sediment, surface water and groundwater at a high quality level, and to improve capacity of the WRI to monitor the quality of surface and groundwater.

3 Methodology

Water quality is based on the end use to which it will be put. It is on this basis that it is analysed for its physical, chemical and biological composition. To determine water quality, scientists first have to make a trip to the water source to obtain a sample. This sample is then analysed later on in the laboratory for such characteristics such as dissolved mineral content, acidity (pH level), dissolved oxygen content, turbidity and bacterial concentration. Some of these parameters can be determined at the stream. In developing countries, the monitoring process is entirely manual. The measured parameters are then compared to numeric standards and guidelines to determine if the water is suitable for its intended use. Standards and guidelines are established to protect water for designated uses such as drinking, recreation, or agricultural irrigation. Standards for drinking water quality ensure that public drinking water supplies are as safe as possible. The field trip to obtain the sample and the subsequent return trip to analyse the sample in the laboratory often causes considerable delay in the monitoring process and hence may provide inaccurate results due to the high temporal variability of the water. In many rural communities in Africa however, the water is not even monitored at all. In fact, the authorities do not have accurate information on water sources for rural communities. This is mainly due to the cost of carrying out the surveys. Some information does however exist on sources of secured or safe drinking water. The limited monitoring that is done is restricted to water sources in urban areas which in many instances are secure or safe. This proposal is aimed at improving the quality of the monitoring process and inexpensively extending the monitoring process to rural communities. Real-time monitoring would enable early warning to enable appropriate action to be taken. The system being developed can be extended to urban water supplies and used as an early warning system. The development of inexpensive, reliable and robust sensors for water quality measurement is a major design challenge.

Teams of researchers will travel across the selected areas to identify and accurately record the location and nature of each water source used by the local communities. The work will call for the design and installation of sensor systems that will monitor water quality effectively at these locations. These sensors will be able to operate for extended periods of time (3–5 years) without any intervention. The data collected from these sensors will be transmitted wirelessly using the existing telecommunications infrastructure of the country and BTS system onto the server which will be linked to the database where each sensor will have its unique location identifier. A text messaging system will be incorporated into the design to enable alert signals to be notified to appropriate person. The outcome of this work will be a system designed to enable the authorities to be alerted automatically of contamination in drinking water source.

The expected result from this research project will be a comprehensive water quality monitoring framework that will have the following attributes:

- *Innovative*: The framework being proposed will be based on novel approaches in fibre optic sensors (FOSs) developments together with the latest advances in

wireless communications technology, with security being an integral part of the design philosophy.

- *High temporal granularity*: It will now be possible to obtain much higher temporal granularity in systems based on wireless sensor technologies. The system will be able to deliver measurements with a very high time resolution with the possibility of having variable monitoring frequency. The major challenge in this respect will be power requirements.
- *Multi-parameter solution*: A key feature of FOSs is that they can be configured to measure several parameters simultaneously, with low power consumption and no electrical pathways within the host structure. In the proposed framework, each sensor node will be capable of monitoring multiple parameters thus decreasing by a significant amount the size of the sensor node.
- *Data integrity and Presentation*: Most often systems are designed and built without consciously considering security issues. As such any security measures are an afterthought which does not fully integrate with the design philosophy. It will be important to ensure that secure protocols are designed and implemented for the data transmission and for the physical sensor units. Data obtained must be kept on a secure server and be presented to users in both tabular and graphical formats.
- *Real-time data collection*: Data will be collected and transmitted in real-time with built-in alarm systems. SMS messages will be sent to registered mobile phones to warn of critical changes in the monitoring environment. This will have implications for power consumption and as such needs to be carefully designed.
- *Cost effectiveness*: Compared to traditional systems the system developed will be cheap and cost effective. It will now be possible to obtain all the required information without the associated costs of sending teams to the monitoring sites.
- *Scalability*: The system can be developed as a framework for a nationwide monitoring system as such it will be easily scalable.
- *Reliability*: A key feature of the proposed system will be its reliability and ability to operate unmonitored over extended periods of time.

The parameters to be tested may include the most important amongst the range of key parameters: mercury, cyanide, lead, arsenic, pH level, temperature and dissolved oxygen content. Other parameters of interest are total coliform and *E. coli* as well as other agents of bacterial and viral infection. For small water bodies, bacterial contaminations in the form of faecal and total coliform contamination are the most important [6, 8].

Government and local authorities often require a reliable information system which they can consult to know the state of play at any given time. It would enable them to take effective action to address emergencies. In addition, the system will feature an alert system that will warn users of any critical changes in the monitoring environment. The following target groups have been identified as the potential beneficiaries of such a proposed system.

- Rural communities: It is known that over 80 % of people without access to safe drinking water live in rural communities. The system will incorporate an alert

system where an SMS text message will be sent to the telephone number of a local leader to warn of a deteriorating water condition.

- Local Authorities: More often than not, government agencies have been accused of not responding appropriately to emergencies such as an outbreak of cholera. The fact of the matter however is that these agencies simply obtain the relevant information often too late for any effective action to be taken. Government agencies therefore at local, regional and national level will greatly benefit from this project as it will provide a cost effective monitoring and surveillance system for them.
- Policy makers: This project will give policy makers the necessary tool to formulate policies that are relevant to the water needs of people in rural communities.
- Research Organisations/Researchers/Academia: For this target group, the sheer amount of data points generated by the project would be like a gold mine. It would enable them to understand more clearly water pollution and its causes.
- Participation: The project must involve close collaboration with relevant target groups.

4 System Architecture

A number of WSNs have been proposed for water monitoring [4, 5, 10, 11, 14]. Most of these are based on solid state electronics sensors and measure only the most basic of parameters, temperature, dissolved oxygen, conductivity, pH, turbidity and water level. To meet this requirement of monitoring these various parameters a number of off the shelf solid state electronic sensors are used. This inevitable adds to the cost, size and power consumption of such systems. Biofouling has been identified as a major problem in these systems, where bacteria are attached to the sensor nodes and thus interfering with the accuracy of the readings. A good review of trends in remote monitoring of the physical environment with particular focus on water bodies, mainly in the United States, is provided in [5].

The diagram of the system being proposed is shown in Fig. 1. It consists of a monitoring node, a database station and a remote monitoring station. Unlike a traditional network, there is no requirement for the nodes to communicate with each other. Each node can communicate directly with the database station. The nodes will independently perform data acquisition by sensing the physical parameters of its environment and if required perform the processing of the raw data locally to obtain the required information. Performing local processing however does add up to the power budget. It is therefore desirable to do this at a central office. At the heart of the system will be FOSs. Most sensors are designed to measure a particular physical parameter known as the measurand at a specific location using a sensor with sensitized tip. FOSs offer the advantage of not suffering from electromagnetic interference and having high bandwidth, compactness and low power, and they can be produced cheaply using components developed from the telecommunications industry. These sensors can be configured to measure several

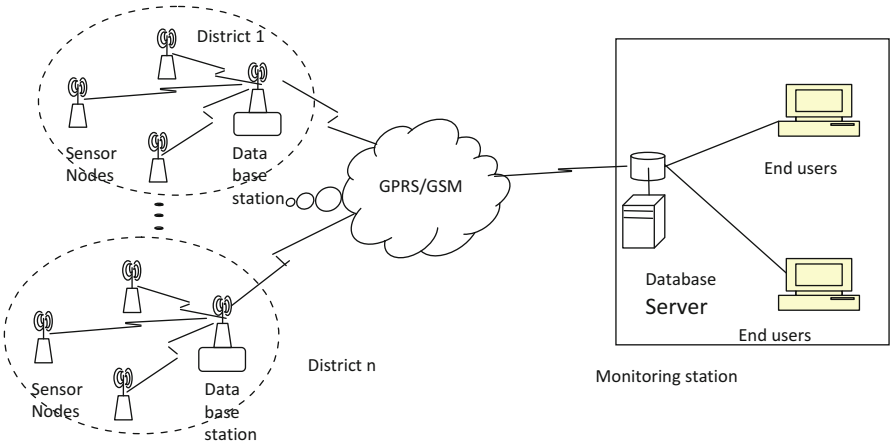

Fig. 1 System architecture

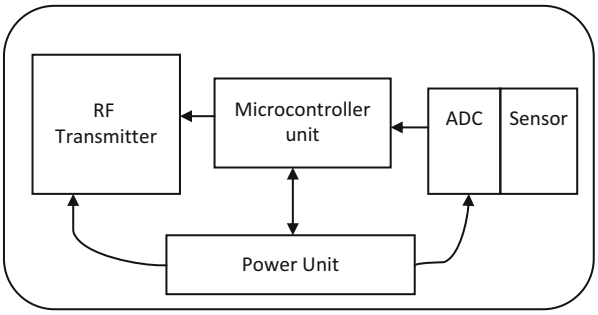

Fig. 2 A generalized schematic diagram of a typical Sensor node

parameters simultaneously thus making them even more compact and also reducing power consumption. Because of their compactness and design, it is possible to measure several physical parameters and multiplex the data from several sensors on a single node.

4.1 Power Requirements

The outline design of the hardware for the sensor node is shown in Fig. 2. It consists of the sensing unit, the transmitter, a microcontroller and the power unit. Realistic and cost effective sensor networks can only be built if the power consumption at the node is low, typically less than 100 μW of average power, low operating and system cost and small size. Being able to minimize the power consumption of the nodes will ensure long lifetime. Obtaining low power consumption is therefore

a major challenge in large scale deployments. Ideally the nodes should be self sufficient in meeting their power requirements, i.e. do not require changes of the power source during their operational lifetime. A number of energy sources are available for use in such deployments. Lithium batteries are easily available and cheap and can provide power densities of up to 100 $\mu W/cm^3$. The operational lifetime of these batteries can be up to 1 year. Importantly batteries do not depend on their environment for their operation even though environmental conditions may affect their performance. The crucial parameters, such as voltage, energy density, specific energy, self discharge, of different types of batteries are provided and reviewed in [14] to help a systems designer in reaching a decision on what is most suitable for their purpose. Solar cells have become popular of late; however they suffer from low conversion efficiency. Reported efficiency rates are in the region of 18–24 %. They can however provide very long lifetimes and power densities are in the range of 10 $\mu W/cm^2$ for outdoor operations to about 15 mW/cm^2 for indoor operation [3]. Their operational performance unlike batteries depends very much on environmental conditions with best performance being achieved under strong sunlight. Apart from solar cells, energy can be harvested through other means, such as mechanical energy from vibrational converters, thermal energy from heating furnaces and sources, electromagnetic energy from inductors and transformers. Other sources are wind and fluid energy which results from the flow of liquids and wind. In rural communities in Africa the ideal energy source might be solar cells used in combination with batteries, charging the batteries by day and using the stored energy of the battery at night.

A typical sensor node performs three basic activities, sense the physical environment, computation such as data compression and communication of the acquired information. In the power budget analysis for sensor nodes, most power is expended on communication, thus it is important that an energy efficient transmitter is used. When active, the transmitter can consume over 99 % of the budget power. To conserve energy, it is important to employ an efficient power management strategy. For the framework under consideration, the transmitter might be in a persistent "sleep" mode only waking up when required to perform the transmission. Using such a strategy, the authors of [8] calculated the energy utilization of a Tyndall transmitter mote. In Stojcev et al. [14] detailed information on power consumption of a good number of microcontrollers and radio transmitters is provided. The current consumption of radio modules typically range from under 10 mA in lower power device to over 50 mA for high power devices. Table 1 shows the main parameters of some typical radio modules available on the market.

There is a wide range of microcontrollers available on the market and Table 2 gives the power consumption of some of them. A suitable microcontroller will be selected from a range of commercially available sources.

From the information provided in [14], a transmitter module can be selected to meet the specified requirements of the application, such as low power consumption and low voltage and ability to operate in the industrial scientific and medical band. In order to meet the low power requirements we will need to consider use of the RF transmitter with good transmitting range up to 2 km. The power unit will be built

Table 1 Main parameters of commercially available transceivers

	CC2420	MRF24J40MB	CYRF6936	AMB8310-EM	MICRF620
Current consumption Rx	18.8 mA	25 mA	−97 dBm	27 mA	12.0 mA
Current consumption Tx	17.4 mA	130 mA	+4 dBm	96 mA	24 mA
Sleep	20 μA	5 μA	1 μA		280 μA
Voltage	3.6 V	3.3 V	3.6 V	3.6 V	2.0–2.5 V
Data rate	250 kbps	250 kbps	DSSS (250 kbps), GFSK (1 Mbps)	57.6 kbps	<20 kbps
Operating frequency	2.4 GHz	2.4 GHz	2.4 GHz	868.00–869.70 MHz	430–440 MHz
Range		1.2 km	50 m	2 km (10 kbps)	
Modulation scheme	DSSS		DSSS, GFSK	FSK, GFSK, ASK/OOK	FSK

Table 2 Main parameters of some typical microcontrollers

	EM6603	STM8S003K3	R5F101FAAFP	STM32F103
Voltage	3.6 V	5.5 V	5.5 V	3.6 V
Current consumption active	1.8 μA	0.58 mA	66 μA/MHz	66 mA
Current consumption sleep	0.1 μA	1.6 mA	0.70 mA	45 mA
Core	32 KHz	16 MHz	32 MHz	72 MHz
Memory	2k × 16 ROM 96 × 4 RAM	8 kb Flash 1 kb RAM	16–512 kB Flash 2–32 kB RAM	256–512 kB Flash 64 kB SRAM
CPU	4 bit	8 bit	16 bit	32 bit

Fig. 3 Data format

from solar panels with matching batteries. The monitoring station will have a main server with a fixed IP address. It will include a database and monitoring software. It will feature the Microsoft windows server 2003 operating system or above and latest professional edition of Microsoft SQL server or appropriate equivalent database management software. Additional hardware including servers and user end PCs with appropriate specifications will be required.

4.2 System Software

Many WSN applications are designed on a push/pull basis; the sensor nodes transmit data at set times but can also be controlled and made to respond to events. In our proposal, the nodes will initially only be sending out data at specified time intervals. The operating system to run on the gateway device will be the open source TinyOS. This has a network protocol and a distributed server. This operating system has flexibility in scheduling. The structure of the data format that will be sent from each sensor will be as shown in Fig. 3. It is made up of the sensor ID which uniquely identifies the sensor that is mapped to particular locations. The DATAID will contain information on the type of data, e.g. temperature, pH level, dissolved oxygen content etc. The TIMESTAMP gives information on the date and time the data was acquired. Then the ACTUALDATA gives us the value of the data. Finally a check code is included for the purpose of error detection.

4.3 Methods of Implementation

A project of this size and complexity would involve many institutions and organizations and requires good planning from the onset to ensure a successful outcome. The discussion in this section will focus on: Field survey and Sensor system design.

4.4 Field Survey

The first step in the monitoring process is the field survey. This survey will help identify all water sources in selected rural communities. A project team will be sent out into the rural community, and working with the local authorities and community leaders, they will identify and accurately locate all sources of water in that local

community. The team would also use the opportunity to educate the community on simple and effective methods of checking their water and making it safe for domestic use. The team would also acquaint the community with the project aims.

4.5 Sensor System Design

FOS design is now a mature technology. It is now possible to measure nearly all the physical parameters of interest using this technology. There are three main techniques by which measurements are made; the means of sensing this depends on whether one is measuring a change in intensity in a light beam or phase changes and as such are referred to as intensity sensors or interferometric sensors respectively. Intensity sensors depend on either Rayleigh or Raman scattering, attenuation of transmitted light due to absorption etc. Interferometric sensors on the other hand have been demonstrated based on magneto-optic or the laser-Doppler effect amongst many others. The second technique by which measurement can be made is the extent of sensing, does the sensor operate over a single point or over distribution of points. Sensors in this category are therefore classified either as point sensors or distributed sensors. The final technique depends on how the measurands act with respect to the fibre, internally or externally and hence we have either intrinsic or extrinsic sensors.

Optical sensors do offer certain distinct advantages over conventional sensors and these have motivated research and development in this field. These include greater sensitivity, electrical passiveness, freedom from electromagnetic interference, wide dynamic range, point and distributed configurations as well as multiplexing capabilities.

4.6 Training, Education and "spin-off" Aspects

Education, skills and knowledge transfer and technology translation should be integral to this type of project and its success. It is proposed that training workshops are organized as an integral part of this project. The workshops will lay the foundation for the future and provide a close interaction between those involved and stimulate ideas for future training and education.

4.7 Justification of Activities

Water quality management aims to control the physical, chemical and biological characteristics of water and data collection is a key element in that process. The requirements of water quality can only be usefully determined in the context of its suitability for the intended purpose. The first step in traditional water monitoring programmes is to define the objective of the programme. For the data collected to

offer useful information the programme objectives need to be very clearly defined. The data generated from the monitoring activity may be of use at the local, national or even international level. Rational water resource management decisions can only be based on the information obtained from the monitoring programme. This survey will provide the basic background knowledge on existing water quality conditions for each identified water source. This information would be useful to the sensor design team as it will determine the particular type of sensors required for each location. The database design team will also require this information to feed into their system.

4.8 Establishment of an Electronic Database

Some survey work has been done in the past; however these results are not readily available or accessible. In most of the instances the results obtained from those surveys have been paper based. A distinctive characteristic of this project is to create an electronic database of the water sources and the results obtained from the monitoring process. This database will in the first instance contain the raw data on all water sources in specified areas. It will be mapped into a Geographic Information System, to enable users in any part of the world to locate water sources for the specified area. The database will later on be populated with measurement results acquired from each location.

4.9 Sensor Design

A key activity in this project will be the design of the sensors. At the heart of the system will be FOSs to monitor the parameter(s) involved, which have become an established and matured subject finding application in environmental monitoring, civil engineering and other major industries such as petroleum and energy production. Most sensors will be designed to measure a particular physical parameter at a specific location. This will be a challenging task because of the wide range of parameters that may be required to be monitored and will require the development of novel methods in physics. The approach being proposed builds on the expertise and reputation of the work done at City University [1, 2, 9, 13, 15, 20–23], UK in creating novel optical fibre sensor solutions where the “unique selling point” is that (1) FOSs regularly permit measurements that are either impractical or uneconomical with conventional electronic measurement technology and (2) are well suited to the systems integration needed for the applications considered. Typically, FOSs are extremely small in diameter, very light, resistant to corrosion and fatigue, capable of wide bandwidth operation and systems of such sensors can be configured to measure several parameters simultaneously, with low power consumption and no electrical pathways within the host structure. FOSs, including using in-fibre Bragg Grating

technology [9], will be configured for temperature, TSS and turbidity measurement and chemical measurements, e.g. dissolved oxygen and colour will use specialist fibre coatings, changing the level of intensity of a fluorophore incorporated in the coating or an absorption change to signal the change in the measurand level. Layer-by-layer deposition technique has proven very satisfactory method to develop different tailored multilayered thin films to coat the different fibre substrates that may be used. The process of deposition is based on electrostatic attraction and can be used with various reagents such as polymers, nanoparticles, metals, dyes, quantum dots, nanotubes and biomolecules such as enzymes and proteins. In this technique, a charged substrate is alternatively immersed into polyanion and polycation solutions to build up the required multilayer coating. Changing the optical properties of such sensitive films with changing analyte concentration is based on a recognition measurement. For instance, the variation in parameters such as the refractive index, absorbance, reflected optical power, transmitted power and the wavelength at maximum absorbance are several examples of such optical properties that will be exploited. The approach to toxicity and total coliform will be through specially designed biosensitive coatings which change their characteristics in the presence of the contaminant. Molecular imprinting has been extensively demonstrated over the last three decades as a versatile technique for the preparation of synthetic molecular receptors capable of the selective recognition of given target molecules.

4.10 Wireless System Design

Until recently WSN was the preserve of applications in sophisticated or military domains. With continuous advances in technology and reductions in cost, WSNs are finding increasing use in civilian applications. They will play an increasingly important role in the day to day life of people in this century and the scope for doing this now can extend beyond the developed world due to the availability of systems. Autonomous system of networks, comprising sensor nodes can be used to accomplish various monitoring tasks, facilitating the computation of the data and creating and transmitting this information to a local control station. Increasingly these sensor nodes are being defined as large in number, tiny and low cost devices, ideally integrable with low cost sensors.

4.11 Website Design

One of the main visibility actions for this proposal will be the design of the website. This will grant the public and users access to the data collected. The website will be linked dynamically to the database and provide real-time information on quality of water in specified areas. Without this website the information collected will not be readily available and thus is an essential element of the project.

5 Conclusion

Of the more than 700 million people worldwide who need access to an improved source of drinking water if the MDG are to be met, more than a third live in sub-Saharan Africa. Accelerated progress is thus required in this area in order to achieve the target. It is estimated that over 4,100 children die every day due to dirty water. Community water supplies, especially in developing countries, are more frequently associated with outbreaks of disease than are urban supplies. Half the people in the developing world suffer from diseases associated with inadequate water supply at any one time and half of hospital beds worldwide are filled with people suffering from water related diseases. Huge disparities exist in access to safe drinking water. In rural areas of sub-Saharan Africa it is as low as 5 %. Investing in creating a reliable source of potable water is a key aim for improving the well-being of the population of Africa: it will reduce waterborne disease outbreaks and the suffering and costs to the economy associated with that. Diarrhoea is responsible for almost 1.5 million deaths each year and mainly of young children. It is the third largest cause of death from infectious disease. According to the WHO, 90 % of those deaths could be prevented through safe water. Progress in providing safe sources of drinking water for these communities has been slow to date but the proposed framework aims to take a significant step forward to improve the situation through new research and associated activity. The work will tackle a major public health issue, doing so by developing new instrumentation and associated techniques, while also providing high quality education and training.

Acknowledgement F.A. Katsriku gratefully acknowledges the support of CAPREX and Alborada Research Fund during the period of completing this work.

References

1. L. Alwis, T. Sun, K.T.V. Grattan, Optical fibre-based sensor technology for humidity and moisture measurement: review of recent progress. Measurement **46**(10), 4052–4074 (2013)
2. J. Cao, T. Sun, K.T.V. Grattan, Gold nanorod-based localized surface plasmon resonance biosensors: a review. Sens. Actuators B **195**, 332–351 (2014)
3. Y.H. Chee, Ultra Low Power Transmitters for Wireless Sensor Networks. Ph.D. thesis, University of California, Berkeley, 2006
4. T.L. Dinh et al., Design and deployment of a remote robust sensor network: experiences from an outdoor water quality monitoring network, in *32nd IEEE Conference on Local Computer Networks LCN*, IEEE, 2007, pp. 799–806
5. H.B. Glasgow et al., Real-time remote monitoring of water quality: a review of current applications, and advancements in sensor, telemetry, and computing technologies. J. Exp. Mar. Biol. Ecol. **300**(2004), 409–448 (2004)
6. R. Henderson et al., Fluorescence as a potential monitoring tool for recycled water systems: a review. Water Res. **43**, 863–881 (2009)
7. B. Jamie, R. Balance, *Water Quality Monitoring – A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes* (UNEP/WHO, 1996). ISBN 0 419 22320 7

8. K.G. Lamka et al., Bacterial contamination of drinking water supplies in a modern rural neighborhood. Appl. Environ. Microbiol. **39**, 734–738 (1980)
9. M.R. Mokhtar, T. Sun, K.T.V. Grattan, Bragg grating packages with nonuniform dimensions for strain and temperature sensing. IEEE Sens. J. **12**(1), 139–144 (2012)
10. B. O'Flynn et al., SmartCoast: a wireless sensor network for water quality monitoring, in *Proceedings of the 2nd IEEE International Workshop on Practical Issues in Building Sensor Network Apps*, Dublin, 15–18 October 2007, pp. 815–816
11. B. O'Flynn et al., Experiences and recommendations in deploying a real-time, water quality monitoring system. Meas. Sci. Technol. **21**, 124004 (2010)
12. S.P. Oliver et al., Foodborne pathogens, mastitis, milk quality, and dairy food safety, in *NMC Annual Meeting Proceedings*, 2005
13. N. Raoufi et al., Fiber optic pH sensor using optimized layer-by-layer coating approach. IEEE Sens. J. **14**(1), 47–54 (2013). doi:10.1109/JSEN.2013.2280283
14. M.K. Stojcev et al., Power management and energy harvesting techniques for wireless sensor nodes, in *Proceedings of the 9th International Conference on Telecommunications in Modern Satellite, Cable, and Broadcasting Services*, 2009, pp. 65–72
15. M.H. Tu, T. Sun, K.T.V. Grattan, LSPR optical fibre sensors based on hollow gold nanostructures. Sens. Actuators B **191**, 37–44 (2014)
16. United Nations Development Programme. We can end poverty 2015, MDG. Available from: http://www.un.org/millenniumgoals/environ.shtml. Accessed 19 Nov 2012
17. United Nations Environment Programme. Global environmental monitoring system. Available from: http://www.unep.org/gemswater/. Accessed 19 Nov 2012
18. World Health Organisation. Water sanitation health [WWW] WHO. Available from: http://www.who.int/water_sanitation_health/facts_figures/en/index.html. Accessed 19 Nov 2012
19. World Health Organization and UNICEF, *Meeting the MDG Drinking Water and Sanitation Target: The Urban and Rural Challenge of the Decade* (WHO, Geneva, 2006). ISBN 978 92 4 156325 3
20. S.P. Wren et al., Preparation of novel optical fibre-based Cocaine sensors using a molecular imprinted polymer approach. Sens. Actuators B **193**(2014), 35–41 (2014)
21. S. Yang, B. Zhou, T. Sun, K.T.V. Grattan, A novel optical sensor platform designed for wireless sensor networks. J. Phys. **450**(1) (2013)
22. B. Zhou, S. Yang, T.H. Nguyen, T. Sun, K.T.V. Grattan, Wireless sensor network platform for intrinsic optical fiber pH sensors. IEEE Sens. J. **14**(4), 1313–1320 (2014)
23. B. Zhou, S. Yang, T. Sun, K.T.V. Grattan, A mobile wireless sensor network platform for use with optical fibre sensors, in *Proceedings of SPIE – The International Society for Optical Engineering*, 2013, p. 8794

Green and Distributed Architecture for Managing Big Data of Biodiversity

Idrissa Sarr, Hubert Naacke, Ndiouma Bame, Ibrahima Gueye, and Samba Ndiaye

Abstract The biodiversity term refers to the totality of genes, species, and ecosystems of a region or the globe. Biodiversity's impact on the human health and the ecosystem is without a doubt very significative. Therefore, the conservation of the biodiversity is becoming an international political and scientific issue since it may have a drawback on climate and the human health or survival. For a sustainable development perspective, several ongoing studies are conducted to analyze, predict, and face biodiversity changes. Such studies require a huge volume of data collected, stored, shared, and exploited intensively by researchers through the world by using web technologies and information systems as GEOBON, LifeWacth, GBIF, MosquitoMap. These systems handle an important amount of computing and database resources that must be optimized for avoiding maintaining useless resources while reducing considerably the energy usage. Actually, the goal of such optimization that we propose in this chapter is to adapt (increase or decrease) the number of resources for dealing with data of biodiversity based on the current load (or number of requests) while ensuring good performances. The benefits of doing so are manifold. First, it fits perfectly with the objectives of green computing or green IT that suggest to define computing systems efficiently and effectively with minimal or no impact on the environment. Second, it is well suited for African developing countries that encounter frequently energy problems and that miss enough funds to maintain complex infrastructures.

I. Sarr (✉) • N. Bame • I. Gueye • S. Ndiaye
Université Cheikh Anta Diop, Fann - BP 5005, Dakar, Senegal
e-mail: Idrissa.Sarr@ucad.edu.sn; Ndiouma.Bame@ucad.edu.sn; Ibrahima.Gueye@ucad.edu.sn; Samba.Ndiaye@ucad.edu.sn

H. Naacke
Sorbonne Universités, UPMC Univ Paris 06, LIP6 Paris, France
e-mail: Hubert.Naacke@lip6.fr

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_2

1 Introduction

The biodiversity term refers to genetic variation, species variation, or ecosystem variation within a geographical area. It may have a great impact on the human health and the ecosystem. Therefore, the conservation of the biodiversity becomes an international and political issue in the last two decades. The main perspective behind this increasing attention is due to the fact that the biodiversity conservation is linked to sustainable development around the world. In this respect, a convention was entered on December 1993 between several countries to face challenges of the biodiversity conservation [18]. These challenges are manifold and are related to a various kind of aspects such as economic, health, science, and politic. A key issue when facing such challenges is to build and share a comprehensive inventory of all species of biodiversity in the world. The motivation of doing so is to follow up species, to understand their life conditions, and to forecast the growth and/or the depletion of their number.

However, these challenges require infrastructures and funds for defining and supporting good policies in order to collect and promote the sharing of the biodiversity data. Data may come from many institutions around the world with different structures and must be shared through the web. In this respect, a set of web platforms are built as GEOBON [10], LifeWacth [13], Global Biodiversity Information Facility (GBIF) [9], MosquitoMap [15]. These platforms bring together the diverse, stand-alone observation instruments and systems that track genetic resources, species, and ecosystems.

Furthermore, such platforms must remain available at any time while ensuring acceptable response time. In this respect, the architecture should be distributed and scalable, which requires powerful infrastructures and several resources. However, the resources of most of the existing platforms are set in a static (or by anticipation) way even though the workload fluctuates over time. In fact, the workload varies in such a way that there are some periods of high overall activity and some other more quiet periods. Hence, a static allocation may lead to a wastage since computational resources may be underused during quiet periods while they keep consuming power. Consequently, the static allocation by anticipation strategy does not permit to cope perfectly with the biodiversity convention that encourages the reduction of useless energy consumption.

The main goal of this chapter is to propose an elastic solution that has the ability to acquire and release resources on-demand in response to workload whose requirements fluctuate over time. The designed approach relies on the data access patterns and the characteristics of the GBIF web portal. Moreover, the proposed solution should be implemented on top of a cloud-based infrastructure that affords computing and storage capabilities with a low cost. The benefits of doing so are multiple. First, it fits perfectly with the objectives of green computing that suggests to define computing systems efficiently and effectively with minimal or no impact on the environment. Second, it is well suited for African developing countries that encounter frequently energy problems and that miss enough funds to maintain complex infrastructures.

Due to the large public we want to reach, we do not dive in deep into the technical aspects of the solution we describe. We refer readers to our work presented in [12] for details about the technical aspects. The rest of this paper is organized as follows. We portray in Sect. 2 the biodiversity and its conservation within African countries. We point out the challenges to face in such a context and we propose some solutions. In Sect. 3, we describe data of biodiversity of the GBIF and some of their use cases. Section 4 presents a green solution to deal with data biodiversity by defining first the data access pattern and highlighting the gain to be had when managing workload in a efficient way. We describe a solution for implementing a green management approach to deal with biodiversity workload in Sect. 5. A discussion of the overall advantages of the approach is given in Sect. 6 while Sect. 8 concludes.

2 Biodiversity in the Realm of African Countries

The conservation of biodiversity is considered as one of the big challenges and key issues of sustainable development since the Earth Summit held in Rio de Janeiro from 3 to 14 June 1992. Actually, the Convention on Biological Diversity as known as the Biodiversity Convention was entered on December 1993 with three main goals: (1) the conservation of the biodiversity; (2) a sustainable use of its components; and (3) a fair and equitable sharing of benefits arising from genetic resources [18]. Even though the ecosystems, species, and genes must be used for the benefit of humans, it is worth noting that natural resources are not infinite and require sustainable use. Hence, the usage of resources should be done without leading to a long-term decline or a dearth of biological diversity.

The specifications and requirements of the convention arouse many issues among which we point out: (1) the sharing of the results obtained from research and development related to genetic resources; (2) the coordination of a global directory of taxonomic expertise; (3) the education and public awareness; (4) the provision of financial resources; (5) the technical and scientific cooperation. These issues show all the complexity of conserving the biodiversity stability that requires several collaborations between scientifics, economists, politicians, sociologists, and so forth. In this respect, the issue of preserving biodiversity is attracting more and more interest, and mainly, in developing countries. Basically, such interests are translated by new policies and structures defined by African government to cope with the goals of the convention. For instance, it is quite impossible nowadays to see an African government without a ministry of Ecology or a ministry of sustainable development. Nevertheless, African countries still face problems to carry out with efficiency and success their policies due to many challenges that we highlight in the next subsection.

2.1 Challenges in African Developing Countries

African countries experiment recurrent problems related to biodiversity because of the fact that many genetic resources are overexploited. For instance, the shortage of good policies of regulation or appropriate tools to control the marine biodiversity leads to an overfishing and a depletion of fisheries that destroys marine mammals and entire ecosystem. As a consequence, illegal, unreported, and unregulated fishing is increasing in developing world since fishermen seek to avoid stricter rules in many places in response to shrinking catches and declining fish stocks. Moreover, in African countries, people partake to poaching for commercial gain, home consumption, and to face the lack of employment opportunities. The body parts of some animals are also in demand for traditional medicine and ceremonies. Furthermore, the dearth of water or rain in many sub-Saharan countries accelerates the disappearance of species and push people to move from an area to another. Such practices and human behaviors speed up the defaunation of forests, the reduction of animal populations, the emergence of zoonotic diseases, such as Ebola Virus, caused by transmission of highly variable retrovirus chains, and so on.

Therefore, it is obvious to observe that dealing with problems of biodiversity in Africa is challenging at many points. Hereafter, a short list of challenges we want to highlight in the context of this chapter.

- One of the main challenges is to perform a reliable and comprehensive inventory of all species biodiversity in Africa. The motivation of doing so is to follow-up species and to understand their environmental and life conditions. Such inventories can be used for predicting and modelling dynamic of species and will help to forecast the growth and/or the depletion of their number. Plus, it may monitor the utilization of genetic resources after they leave a country including by designating effective checkpoints at any stage of the value-chain. However, this inventory requires infrastructures and storage support for long term conservation and for a wide sharing. Such infrastructures are infrequent due to their expensiveness and if ever they exist, they are not enough efficient.
- Another challenge consist in defining strong, fair, and non-arbitrary rules procedures for regulating use of genetic resources or for protecting forest areas. Such procedures must be established for education and public awareness and for prior informed consent and mutually agreed terms. The problem of getting this goal is caused by the language barrier and the shortage of communication support for sensitizing every one of the drawback of overexploiting genetic resources. In fact, people of African countries do not speak the same language, and thus, each message should be translated in various dialects to reach people in remote rural areas. Moreover, some people resist for any change of their traditions or culture. Hence, such African realities make quite impossible or difficult to establish any consent and mutually agreed terms. Nevertheless, new technologies may help to communicate with such people through video messages and Internet. However, Internet and electricity are not present everywhere in Africa and remain entire issues.

- Last but not the least, African governments should create conditions to promote and encourage research contributing to biodiversity conservation and sustainable use. To this end, the creation of worldwide databases for sharing taxonomic expertise and semantic is a paramount issue. However, there is an increasing lack of policies or funds that do not encourage researchers to collect and share taxonomic. Even if some researchers keep working on biodiversity conversation by modelling the ecosystem and predicting its evolution, governments do not valorize their works by promoting them in a large public or rewarding them. Thus, researchers tend to go abroad where they may get infrastructures and supports to develop their ideas.

2.2 Solutions and Contributions

One may see that these challenges cited above cannot be faced without infrastructures and funds for defining and supporting good policies in order to collect and promote the sharing of the biodiversity data. Data may come from many institutions around the world with different structures and must be shared through the web. To this end, a set of web platforms or web portals are devised in developed countries such as GEOBON [10], LifeWacth [13], GBIF [9], MosquitoMap [15]. These platforms bring together the diverse, stand-alone observation instruments and systems that track genetic resources, species, and ecosystems. Moreover, these platforms can integrate biodiversity data with data on climate and other key parameters in order to fill gaps in taxonomic and biological information and speed up the pace at which information is collected and disseminated.

Since such platforms are designed for public and free use, thus, they can be used by African researchers for uploading or downloading data of biodiversity related to the continent and their research. In other words, these platforms can be an alternative solution of the challenges we cited in the previous section. Actually, African countries can rely to these information systems to face challenges caused by shortage funds required to build and maintain complex infrastructure for biodiversity data management.

Furthermore, such platforms must remain available at any time while ensuring acceptable response time. In this respect, the architecture should be distributed and scalable, which requires powerful infrastructures and several resources. However, the allocation of resources or their design is often done based on the prediction of maximum needs that the system could face in terms of computation (maximum peak load or simultaneous user requests) or in terms of storage (maximum volume of data). With the storage perspective issue, it is worth noting that resources are and remain generally appropriate and their extension are pretty simple. However, the allocation by anticipation of computational resources may lead to wastage since the appearance of the maximum load is infrequent. Therefore, allocated computational resources are often idle while they keep consuming power for cooling and requiring human resources for administrative tasks.

Consequently, the allocation by anticipation strategy does not permit to cope perfectly with the biodiversity convention that encourages the reduction of useless energy consumption. To deal with this issue, we propose an elastic solution that has the ability to acquire and release resources on-demand in response to workloads whose requirements fluctuate over time. The motivation of doing so is to optimize and/or reduce the number of resources in a dynamic fashion and to be able to get the objectives of "green computing" or "green IT."

Furthermore, African researchers need enough and sophisticated local infrastructures to store data they download or on which they want to work on. However, such infrastructures are not always available due to shortage of funds and lack of expertise. Hence, a solution to this problem is using a Cloud-based infrastructure to afford computing and storage capabilities with a low cost. In fact, Cloud computing allows companies to avoid upfront infrastructure costs, and focus on their businesses instead of infrastructure. Moreover, Cloud resources are usually shared by several users and are also dynamically adjusted per demand to meet fluctuating and unpredictable business needs. The Cloud has the advantage to maximize the use of computing powers thus reducing environmental damage since less power, air conditioning, rackspace, and so on is required for a set of functions. Plus, multiple users can access a single server to retrieve and update their data without purchasing licenses for different applications. This strategy fits well in the context of African countries where researchers have not enough financial support to bring into their team all skills or applications they need.

Briefly, the main contribution of this chapter is a combination of two mechanisms, namely, a cloud-based infrastructure and an elastic data management strategy for biodiversity. We apply our solution on the GBIF information system that we describe in next section before describing how we manage data of biodiversity in a green fashion.

3 Describing Data of Biodiversity: GBIF Case

The GBIF is an international organization created in 2001 to ease the collect, integration, and share of primary biodiversity data [9]. The GBIF database is hosted in Dannark. Data come from many institutions from around the world and are related to plants, animals, fungi, and microbes of the world. That is, African countries are invited to invest in such infrastructures by uploading data related to species they discover and identify. The GBIF mainly focuses on making scientific data on biodiversity available via the Internet using web services and can be used as a support of communication between politicians, policy makers, researchers, and the general public together. That is, the GBIF mission is to facilitate free and open access to biodiversity data worldwide for building and catalyzing sustainable development. Another goal of the GBIF is to promote participation and working through partners for collecting biodiversity data, building a computing and storage architecture to allow their integration, designing protocols, and standards to

insure scientific integrity and interoperability, in order to make analytical tools for improving decision-making. Actually, the GBIF objectives can be summarized as follow:

- an international project for inquiring the global primary biodiversity data (from genes to ecosystems);
- an information architecture that makes biodiversity data accessible and searchable through a single portal;
- a network of international organizations that play the role of information assistance and training for data providers and users;
- a reliable scientific data on which scientific analysis can be processed in order to establish trends for biodiversity management.

3.1 Data Type and Data Model

GBIF portal gathers several types of data, namely, primary data and metadata. Basically, primary data refer to species information (e.g., taxonomy and owner), the details of their observations (e.g., geographical position, country, and region) and documentation in the form of audio-visual (e.g., photographs, videos, and audio). However, metadata describes the primary data as well as specifying the details of their providers, their collections, and so forth.

Moreover, the database is made of collection that contains many occurrences of a species. An occurrence describes characteristics of a specimen and contains many taxonomic and geographical position fields. In addition, an occurrence represents the basic unit of information in the database. It is worth noting that a collection belongs to a provider and linked to a country. Furthermore, the database contains more than 400 million hits (occurrences) from more than 10,000 datasets and 560 providers. Data are structured by using a relational data model and stored in such a way that all the database is mapped into two kind of tables: a main relation that contains all information about occurrences, and a set of side tables (small sizes) containing essentially the metadata of occurrences. Finally, the number of data records available in the web portal increases day in and day out. For instance, from December 2008 to September 2013, the amount of data records has varied from 163 to 416 millions. Hence, this huge amount of resources on primary data of biodiversity data keeps attracting more users to the GBIF portal and leads to an increase of more complex workload.

3.2 Use Cases of GBIF Data

Use cases of biodiversity data are manifold, various, and involve a range of activities. A list of use cases of biodiversity data are available on the GBIF portal [9]. These use cases are set for analyzing the interaction between species,

the evolution of species, extinction risk, socioeconomic importance, the impacts of climate change, etc. In this chapter, we describe use cases of modelling the ecological niche, species co-occurrence, and data input.

Modelling Ecological Niche

Modelling the ecological niche of a species is a process of building a function or a model that uses environmental parameters in order to predict the probability of presence of a given species. This modelling aims at studying a species behavior such as its distribution, migration, threat of extinction with respect to environmental factors, and spatiotemporal dimensions. This modelling can be used in agriculture for assessing the impact of bees on a flower, or to figure out new orientations toward a novel kind of peanut seeds, etc. It is also interesting to model ecological niche in the realms of ecology and breeding in order to identify the areas to be protected for the survival of endangered species, and to measure the productivity of a given animal based on resources and coexistence of several other species.

As pointed out earlier, modelling the ecological niche requires a complex data set containing both data related to the studied species and climatic or environmental data of the geographical regions within which studies are conducted. Data related to species describe their characteristics while environmental data portray physical and chemical data such as temperature, precipitation, salinity, solar radiation, and their interactions with other species as the relationship predator/prey. The GBIF portal provides data related to species description and geographical presence while environmental parameters are obtained from other data sources such as BIOCLIM, WorldClim.

Species Co-occurrence

Calculations of co-occurrence are very common in data analysis of biodiversity. In order to model the interactions and/or dependencies between species co-occurrences are required and are used in community ecology. Co-occurrence modelling consists of proving that two or more species coexist in the same area during the same period. For this to be done, the density of two given species must be greater than a minimum value (threshold) within a time window and a space dimension. Basically, the co-occurrence is computed for each cell of a space. To calculate the co-occurrence of several species, the density of species are first determined within each cell before aggregating all densities for the entire space in order to figure out the co-occurrence of a couple of species.

Data Input

To ensure the quality of data available in a collection, updates are necessary. These updates are correction of existing data or insertion of new occurrences. The corrections (change or delete) are performed when errors are noted on the validity of

data records. These errors can impact the consistency of observations and results of analysis (e.g., aquatic species observed in a desert) or the taxonomic structure (e.g., an herbivore baobab). After each prospecting, new instances are collected and are inserted into the database. This involves treatments identification and validation to check the consistency of the field observations and the risk of redundancy with other data records.

Based on the above use cases, the data type, and the data model, it is obvious to note that data of biodiversity may be considered as Big data because of their volume, velocity, and variety. Dealing with Big data is seen as one of the major challenges of this decade in both computer science and information system. Therefore, the management of the biodiversity data stored and shared through the GBIF portal is a paramount issue. As pointed out in Sect. 2, data of biodiversity must be managed in a green fashion in order to fit into the biodiversity convention. We motivate in the next sections an approach for biodiversity data and we afterwards, present the architecture for that purpose. Finally, we describe solutions for facing workload related to biodiversity data.

4 Green Data Management of the Biodiversity Data

Aiming to manage huge biodiversity databases using a network of distributed machines, we face a data placement problem. We state the problem as follows: given a set of machines (each with its own data management capacity in terms of computation and storage), given a workload of data access requests that continuously arrive, we have to assess where (and when) to place the requested data such that all the requests operate in reasonable time, using the minimal number of machines. Knowing all the requests in advance would allow to find a near optimal data placement as well as a schedule to move data when necessary. However, this does not apply in our context where the workload is not entirely known a priori. Conversely, an arbitrary workload that does not reveal any access pattern would not benefit from any clever placement strategy: in such a case a random placement appears to be the best solution. In between these two types of workload (ranging from fully known to fully random) we investigate data placement strategies when the workload exhibits some regular patterns. Hopefully, biodiversity data access presents some specific patterns that we aim to exploit in order to manage data at lower cost (i.e., in a green fashion).

4.1 Data Access Pattern

The main pattern concerns the data popularity that relates to data taxonomy. The users generally target a predefined set of species because they have to decide about an action to preserve, understand, or forecast the evolution of such species.

In many cases, the users enter into a sequence of several requests about the same set of species for a period (from hours to days) that is long enough to deserve dynamic data management adjustment. We have also observed that some users involve several other users that in turn submit their own requests about the same species. We assume that such kind of sudden interest spread (or popularity peak) tends to occur at an increasing rate and an increasing intensity (i.e. higher peak level) as an effect of the increasing socialization level of the users. In other words, the users tend to crystallize around some data for some time, then switch their focus to another data that becomes the next collective target, and so on. The second pattern relies on the observation that the workload globally fluctuates over time. The high number of users does not actually smooth the aggregated workload level. There are still some periods of high overall activity and some other more quiet periods. This can be faced by either a static or dynamic resource provisioning at a global level in order to solve the green data management challenge.

4.2 Drawback of a Static-Based Resources Allocation

In fact, for most of the analysis done over the data biodiversity, a user has to download occurrences of species that are required for a given study. In addition, climatic and environmental data may also be required for a purpose of a study. In such a case, data have to be gathered or integrated through aggregation and join queries that are very expensive, mainly, for large partition of data. That is, the amount of data used in the management of biodiversity are tied to a heavy workload because of the number of potential users.

To handle both data storage and workload issues, the data are partitioned and distributed on different storage nodes. Rather, an important amount of data stored on one node may be related to a few number of species (plant or animal) that are most stressed or required based on their importance or role during a bounded time interval. That is, the workload of the data of the biodiversity fluctuates over time. Thus, even though data are distributed, the system may experiment bottlenecks at some resources. To avoid such bottlenecks appearance, a well-known solution is an over-provisioning approach that allocates a fixed high number of resources. This is done at the design step by allocating resources according to the highest expected load for each partition. In other words, knowing the size and the content of a partition, and how frequently it is accessed, one can estimate the number of resources to set for facing all the workload.

The problems with over-provisioning are that at a given time, the number of the fixed resources may be underutilized, which may lead to a waste of power. Another approach would be to set an average configuration, which consists of allocating resources according to the average load of the system. This approach seems to be better whatsoever, it can lead to a situation in which the allocated resources may not face a peak load higher than the average.

4.3 Benefits of an Elastic-Based Allocation

Using an elastic-based approach is a mechanism of dimensioning resources configuration regarding to the size of the workload. Since workload can be very huge, an elastic solution requires to have on-hands unbounded resources. Unbounded resources can be obtained from a cloud computing infrastructure that holds huge resources.

Cloud computing and/or elastic computing have emerged as successful paradigms for scaling up with a low cost. One of the main factors of this success is that elasticity aims at allowing resources variations in terms of amount while the system works. This strategy gives the possibility of having initially a minimal resources configuration that is increased or tuned, if need be, in order to ensure low latency or good scalability. In the opposite, resources may be reduced if the workload decreases and requires less than current fixed resources. In order to adapt such elasticity in the realm of a biodiversity data management, we need to design a couple of mechanisms that take into account data access pattern and characteristics of the biodiversity.

Many works have been done for studying elasticity with or without a cloud-based platform [4,5,7,11]. Among those solutions, we point out TransElas [11] that adapts the middleware size according to its load, and a solution described in [12] that permits to adapt the database layer size (processing capacities and storage) to the workload variations. The key idea of these works was to organize the database as partitions and to migrate partitions of a overloaded database to an underloaded one in order to ensure a bounded response time while minimizing the resources.

The migration process is conducted without interrupting the system or the current transactions processing. Several works have focused on this specific feature. In [6] the authors propose a pull-on-demand approach in which an index is updated when a migration decision is taken. The goal is to ensure pursuing workload processing either on the migration source or the destination one. Moreover, we can mention Slacker [1], a middleware solution that uses hot Back-Up tools to copy the database while allowing service continuing during this phase. The migration method is based on the available processing capacity in order to minimize the migration impacts response time.

Another dimension is the resources optimization of the data partitioning or replicas placement. Several works have been done to face this issue such as Schism [3] and Sword [16]. These approaches rely on current load level to place replicas. They use a graph partitioning algorithm to find a placement that prevents data distributed transactions. In the graph, vertices are tuples and edges the co-occurrence of tuples that appear in the same transaction. The main objectives of these works are to provide an improved throughput while providing fault tolerance and scalability for distributed OLTP data management systems. However, the power saving that is crucial in biodiversity context is missing.

More recently, the minimization of the migration cost is taken into account in the study [17]. This approach faces the operational costs minimization by adapting in

an elastic manner the system size depending on the tenant behavior. Moreover, the approach considers and quantifies performance degradation during a migration in both the source and destination servers.

In conclusion, elastic computation is gaining popularity over static provisioning computation and allows to regulate resources use in a flexible and extensible way. Therefore, it affords a great opportunity to avoid to unnecessarily tied-up resources for managing data. Moreover, it minimizes wasted storage space and power, and it exhibits low performance overhead, such that it does not lead to a significant longer latency that jeopardizes performance requirements or incurs extra costs. For this to be possible, we must implement elasticity in a highly dynamic and a transparent fashion, such that it hides all details and cope well with biodiversity requirements.

5 Elastic and Green Workload Management

Our goal is to propose an elastic approach to handle biodiversity data. Based on the data access pattern described in Sect. 4, the workload may fluctuate over short periods of time and reach a peak. Since peaks are infrequent, few resources are initially allocated for the data management. Later, when some data tend to be overloaded, additional resources are added and the workload is distributed efficiently over all resources. In the opposite, when the workload decreases significantly, resources are released for keeping only the required amount of them. In terms of data management, adding a resource implies to assign a partition and its related workload to an ad hoc machine. Respectively, releasing a resource implies to retrieve a partition from the machine on which it was stored. Furthermore, in order to figure out data that have to be moved dynamically, we rely on data partitioning and replication mechanisms detailed below.

5.1 Biodiversity Data Partitioning and Replication

We exploit the partitioned nature of the data access pattern (described in Sect. 4), to partition the data. A user querying biodiversity data usually focuses on a small set of species. Users updating (or adding) data also target their access to some species. Therefore, we can logically partition the data according to the species.

More precisely, the species description of biodiversity data is well organized in a hierarchy called a taxonomy. We take the taxonomy into account to define the data partitioning such that partitions match the user queries as much as possible. The taxonomy has several levels. For instance, a family contains several genus, and a genus contains several species. A query accessing all the species of a given genus (or family) allows to define a genus-wide (resp. family-wide) partition. Once the partition is defined, we decide which resource to allocate to each partition, using a data placement strategy. Several partitions may reside at the same machine

as long as the corresponding workload does not exceed the machine capacity. The placement strategy differs depending on two types of access: (i) for write access, we must guaranty consistent data access when several users require common data simultaneously. Thus, we ensure that at any moment, partitions being updated concurrently are disjoint. (ii) For read only access, we may replicate a partition at several places when more resources are needed for accessing that partition fast enough.

Transactions are update operations (insertion, deletion, modification) processed on the GBIF data. We consider two kind of transactions, namely, single-partition and multi-partition transactions based on the number of partitions required by a transaction.

5.2 Handling Single-Partition Transactions

Obviously, a single-partition transaction is accessing only one partition. It consists of inserting either new species or modifying/deleting an existing one.

A straightforward method to process the transactions consistently is executing them in a serial way on a single physical machine. However, when the incoming transactions rate is increasing, a single machine may not be able to process all the transactions in due time. Therefore, we need to migrate some partitions in order to balance the load over several machines. To this end, we continuously track the load of each partition at every machine. We identify the data partitions on which the peaks occur or from which response time is lengthening. We identify the data partitions from which the peaks occur or from which response time is increasing. The migration algorithm operates in order to permit partitions located on the same machine to be used during the migration period.

The migration algorithm is decentralized to ensure scalability. That is, each machine facing a peak is responsible to migrate some of its partitions. A common protocol coordinates the allocation of the available machines that are candidate to accept new partitions.

Illustrative Example

Given a set of data partitions P_1, P_2, P_3, and P_4 stored on a storage resource DB_1 as depicted in Fig. 1. Each partition is represented by a square and we set the color of squares as follows: (1) light gray for an idle partition (no transaction accessing it), (2) black if the partition is heavy required (overloaded) and another color for an intermediate situation (neither idle nor overloaded). One can see that at 8 o'clock, all partitions of DB_1 are idle. However, at 9 o'clock, DB_1 is overloaded because of the partition P_4. In such a situation, the first step after identifying the partition causing the issue is to find where to move it. Consider DB_2 a second storage resource on which we have already three partitions P_5, P_6, and P_7 (see Fig. 2). Since DB_2 is

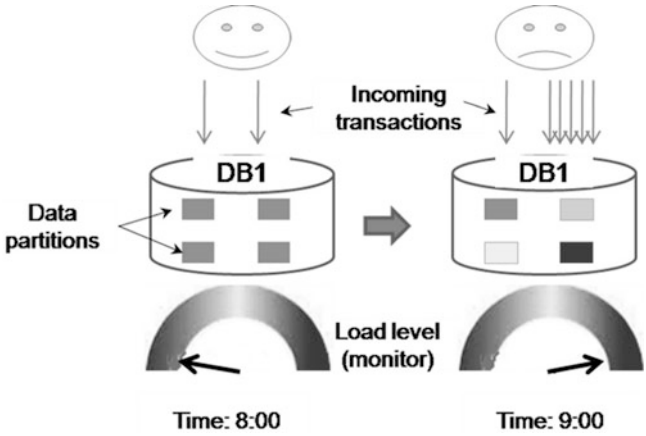

Fig. 1 Load of DB_1 between 8:00 and 9:00

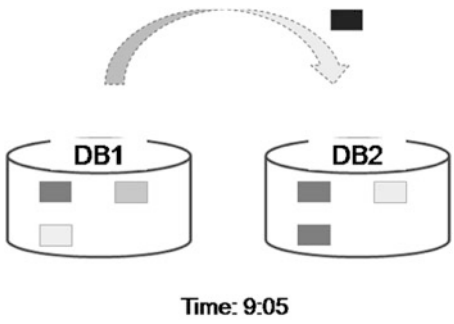

Fig. 2 Migrating P_4 from DB_1 to DB_2

not overloaded (no black square), thus P_4 can be moved on it and all transactions requiring P_4 are also moved on DB_2. It is worth noting that before migrating a partition to another point, we assess if this latter will not be overloaded. That is why in Fig. 3 everything is perfect after the migration completed. Finally, if ever there is no available resource, a new DB instance will be created for receiving P_4 thanks to the virtualization and elasticity capabilities of the cloud infrastructure we use.

5.3 Dealing with Multi-Partition Transactions

A multi-partition transaction accesses at least two partitions, and should be executed on a single physical machine to avoid distributed transactions over a network, which may break scalability performances [19]. Therefore, given a transaction requiring many partitions, we first group the partitions on a single machine if ever they are not already grouped. Precisely, groups are defined in such a way that each transaction is performed on only one group. However, the grouping protocol may lead to gather a high number of partitions into the same group. In such a case, migration, if any, will

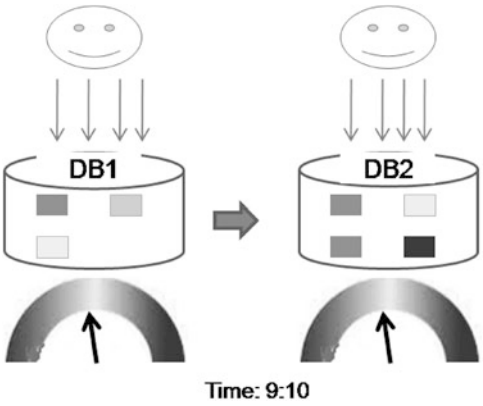

Fig. 3 Load of DB_1 at 9:05

last more time while the related workload can reach rapidly a peak. One solution to avoid large groups is to limit the number of partitions that can be placed into a group.

Moreover, we track the load of each physical machine and migrate groups accordingly. However, when migrating groups of partitions instead of moving partitions one by one, the peaks may occur more frequently and avoid to get a good load balancing. With this in mind, it is important to be able to predict the occurrence of peak load and use this prediction to make the right choice when moving partitions or groups. Such prediction is beyond the scope of this chapter and we recommend reader to have a look on works described in [12], where we propose how using a social graph may help to forecast peak of workload.

6 Discussing About the Overall Gains of Our Solution

In this section, we summarize and discuss about the gains of our approach to manage workload of the biodiversity data. In short, we demonstrate through the paper that using elasticity and a cloud-bases infrastructure has the advantage to reduce cost and resources for managing data biodiversity with a green fashion. Rather, our solution has other advantages in terms of latency, scalability, parallelism, and data placement.

6.1 Controlling Parallelism

Queries for analyzing data of the biodiversity are complex and require heavy computation that may overpasses both the computational and storage capacity of a single machine. Hence, computing such queries must be done in a parallel manner on several machines. However, we do not want to parallelize as much as possible because this would create too much replicas, which require additional overhead for

data consistency control. Therefore, we design an algorithm to control the degree of parallelism for queries. To this end, we rely on a cost model for choosing an optimal plan to process a query. Basically, each query may be subdivided into many sub-queries that can be processed with several plans. The best plan is the one that reduces the amount of data to transfer from a machine to another one. Our algorithm coupled with the grouping strategy we described in the previous section ensure that parallelism is controlled. The motivation of doing so is that transferring data via a cloud infrastructure is costly. That is, reducing such transfers minimizes the financial cost.

6.2 Bounding the Response Time

The response time of a query must be bounded to satisfy the user requirement and depends on the current load status of the computational resources. Hence, to bound the response time, we need to track load of each resources when several users are simultaneously attempting to access them. Afterwards, the less loaded resource in terms of computation capabilities is chosen when a query can be processed by a set of resources. In addition, we record the capabilities of all resources in order to figure out the resource on which the response time is the lowest. The lowest time is estimated based on statistics, current loads, and capacities of available resources, and finally, the data transfer time.

6.3 Optimizing the Data Placement and Replication

Our grouping protocol described above coordinates the replication decisions as well as the data placement of the GBIF data. It prevents to overload resources that have enough and available capacities.

Moreover, replicas are created on the fly for an urgent need that corresponds to face a peak load. Once this load disappeared, the replica is deleted to reduce cost of maintaining consistency. Asynchronous replication is used to avoid lengthening response time. In most of the cases, users can be satisfied with slightly out-of-date replicas. Thus, synchronization of replicas is infrequent and planned for periods of low workload. The advantage of this infrequent synchronization is to reduce data transfers and allow saving money.

7 Related Work

The focus of this work is the management of large volume of data of the biodiversity. Many studies have been conducted to face issues of managing the biodiversity data [9, 14, 15]. Most of the proposed solutions are often oriented to a specific concept

such as a thematic-based approach [15] or a country-based one [2, 8]. For instance, thematic-oriented solutions permit the sharing of data by using thematic. Even though these solutions are frequently used and useful, data are highly tied to a small topics, and therefore, users who want information about other non-linked topics cannot be satisfied. Furthermore, some studies [14, 15] have focused on including analysis on biodiversity data. One of the well-known of such studies is MostiquoMap [18]. It is a very interesting approach that allows users to launch their analysis via a graphical user interface (GUI) from which results are displayed as a map after. However, this approach is tightly tied to MostiquoMap and may not be suited for various kind of biodiversity. Moreover, work described in [9,14,15] uses a centralized database for hosting all data and therefore, all requests are processed in a single point. Such approach has the drawback of jeopardizing the data availability as well as the scalability of the overall system. It is clear that centralization is not the best solution and particularly in a poor context where electricity is not always insured.

Our solution differs to the previous ones at many points. First, it distributes the data to where they are frequently used regardless of their origins and their themes. Users can also access the data regardless of their locations. Hence both scalability and availability are increased. Actually, our solution appears as a complement to previous solutions that do not directly share data sources or computing resources to meet all user needs. Second, our dynamic mechanism for data distribution and query processing ensures the scalability and integration of new features such as the ability to handle more complex queries and conducting analysis on data of biodiversity.

8 Conclusion and Perspectives

This chapter presents an approach to manage data of biodiversity, mainly, how the workload can be processed in a green fashion and to cope with the challenges aroused in the realm of African countries. In fact, data of the biodiversity may be considered as Big data because of their volume, velocity, and variety. Dealing with Big data is seen as one of the major challenges of this decade in both computer science and information system. Moreover, based on the convention of the biodiversity, several collaborations between scientifics, economists, politicians, sociologists, and so forth are required. Among the objectives of such collaborations, we cite the sharing of the results obtained from research and development related to genetic resources and the coordination of a global directory of taxonomic expertise. For this to be possible, expensive storage and computational infrastructures are required and they should be managed in such a way that the response time remains acceptable even though the workload is heavy. However, such infrastructures are not always available due to the dearth of funds and lack of expertise in most of the developing countries. To overcome these limits, we propose an elastic and cloud-based infrastructure solution that affords computing and storage capabilities with a low cost. In one hand, the cloud offers the advantage to optimize the use of

computing powers and multiple users can access a single server to retrieve and update their data without purchasing licenses for different applications. This strategy fits well in the context of African countries where researchers have not enough financial support to bring into their team all skills or applications they need. On the other hand, elastic computation allows to regulate resources use in a flexible and extensible way. Therefore, it affords a great opportunity to avoid to unnecessarily tied-up resources for managing data. Moreover, it minimizes wasted storage space and power, and it exhibits low performance overhead, such that it does not lead to a significant longer latency that jeopardizes performance requirements or incurs extra costs. For this to be possible, we must implement elasticity in a highly dynamic and a transparent fashion, such that it hides all details and cope well with biodiversity requirements. To this end, we propose some algorithms that track loads of resources and once one resource is overloaded, it processes a migration of partitions that cause the problem. The migration is done in such a way that distributed transactions are avoided and the load is well balanced.

Another issue that we do not take into account in this chapter is analytic workload that are made of complex queries and may last for a long while. Even though, analytic workload is crucial, we could not give more details due to a sake of presentation and space. We plan to portray how we face analytical workload in another paper.

References

1. S.K. Barker, Y. Chi, J.H. Moon et al., "Cut me some slack": latency-aware live migration for databases, in *International Conference on Extending Database Technology (EDBT)* (2012), pp. 432–443
2. Canadian BIF (2013), www.cbif.gc.ca
3. C. Curino, E.P.C. Jones, Z. Yang et al., Schism: a workload-driven approach to database replication and partitioning. VLDB Endow. **3**(1–2) 48–57 (2010)
4. C. Curino, E.P.C. Jones, S. Madden, Workload-aware database monitoring and consolidation, in *International Conference on Management of Data (SIGMOD)* (2011), pp. 313–324
5. S. Das, D. Agrawal, A. El Abbadi, ElasTraS: an elastic transactional data store in the cloud, in *International Conference on Hot topics in Cloud Computing* (2009)
6. A.J. Elmore, S. Das, D. Agrawal et al., Zephyr: live migration in shared nothing database for elastic cloud platforms, in *International Conference on Management of Data (SIGMOD)* (2011)
7. U.M. Farooq, R. Lui, A. Aboulnaga et al., Elastic scale-out for partition-based database systems, in *IEEE International Conference on Data Engineering (ICDE)* (2012), pp. 281–288
8. GBIF France (2013), www.gbif.fr
9. GBIF Secretary: GBIF data portal, GBIF web site (2013), data.gbif.org
10. GEOBON Web site (2013), www.earthobservations.org
11. I. Gueye, I. Sarr, H. Naacke, TransElas: elastic transaction monitoring for Web2.0 applications, in *Data Management in Cloud, Grid and P2P Systems* (2012), pp. 1–12
12. I. Gueye, I. Sarr, H. Naacke, Exploiting the social structure of online media to face transient heavy workload. In *The Sixth Intl. Conf. on Advances in Databases, Knowledge, and Data Applications*, IARIA (2014), pp. 51–58
13. LifeWatch Web Site (2013), www.lifewatch.com

14. Map of Life (2013), www.mappinglife.org
15. MosquitoMap (2014), www.mosquitomap.org
16. A. Quamar, K. Ashwin, A. Deshpande, SWORD: scalable workload-aware data placement for transactional workloads, in *International Conference on Extending Database Technology (EDBT)* (2013)
17. J. Schaffner, T. Januschowski, M. Kercher et al., RTP: robust tenant placement for elastic in-memory database clusters, in *International Conference on Management of Data (SIGMOD)* (2013), pp. 773–784
18. The Convention on Biological Diversity (2013), http://www.cbd.int/
19. A. Thomson, T. Diamond, S. Weng et al., Calvin: fast distributed transactions for partitioned database systems, in *SIGMOD* (2012)

Towards an Ontology-Based Framework for Data Integration: Application to Agriculture and Health Domains in Senegal

Moussa Lo, Gaoussou Camara, Cheikh A.T. Niang, Seydina M. Ndiaye, and Ousmane Sall

Abstract Computer-based systems have reached an advanced level in Africa during the past decade. In Senegal, the agriculture and health domain have led the way by initiating number of projects in ontology-based data integration. In this chapter, we are aiming at two main purposes. First, we present the state of the art in data integration in the agriculture and health domains in Senegal. We focus on the two latest projects: SIC-Senegal for agriculture and SIMENS (the National Medical Information System of Senegal) for the health. Second, we describe the methodological evolution of the ontology-based conceptual framework for data integration stemmed from the lessons learned along the two previous experiences. Finally, we will end this chapter by discussing the new challenges in ontology-based approach in supporting information system design in a larger scale such as climate change risk management. We also discuss the linked open data initiative in Senegal through the KOCC (Knowledge, Ontology, Community and Culture) project.

1 Introduction

Developing countries, such as Senegal, live by the rhythm of public policies in key sectors such as Education, Health, Energy, and Agriculture. These areas are accompanied by international funding from bilateral partnerships or obtained through the support of international organizations such as the World Bank or the International

M. Lo (✉) • S.M. Ndiaye
LANI/UFR SAT, UMMISCO, LIRIMA, Université Gaston Berger,
B.P. 234, Saint-Louis, Senegal
e-mail: moussa.lo@ugb.edu.sn

G. Camara
Université Alioune Diop, B.P. 30, Bambey, Senegal

C.A.T. Niang
Université de Tours, 3, Place Jean Jaurès, 41000 Blois CEDEX, France

O. Sall
Université de Thiès, B.P. 967, Thiès, Senegal

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_3

Monetary Fund (IMF). In return, these funding require commitments from recipient countries, to apply the rules of standard management of public policies, that is, rigorous monitoring of various indicators and the rules of transparency in the use of these funds. Thus, it is important for these countries to have visibility on all programs that will result from these public policies, in order to ensure consistent monitoring and evaluation in due course. In addition, the constant need for information is important for each country, regardless of the relationship it may have with its donors, just to ensure proper conduct of its public policies.

In the case of Senegal, two key sectors interested us: Agriculture and Health. These two sectors are of crucial importance for Senegal which has made them as priority. The goals are, on the one hand, to boost employment and ensure food self-sufficiency, and on the other hand, to guarantee the health to everybody, especially vulnerable people such as women and children. To ensure total visibility of these areas, the requirement goes beyond the simple availability of information but rather on knowledge management.

For these two sectors, the need for knowledge management is at two levels. First, we have the knowledge at the macro level that serve as input for the establishment of global indicators in the monitoring of public policies and, second, we have the knowledge available at the operational level that can be used by the domain stakeholders in order to improve their actions.

In this chapter we will see how we are using ontology-based knowledge management and data integration in each of these two cases. First in the field of agriculture, we will see how our approach enables an actor such as the Society for the Development and Exploitation of the Delta Lands (SAED) to better guide its efforts through a better knowledge management. Then we will see how in the field of health, a national health system would provide a better way in organizing data in order to contribute to the enhancement of epidemiological monitoring. Then, we present our work in putting all these experiences together for designing a scalable ontology-based framework which could allow to any industry sharing resources, and improving knowledge management and decision making. The chapter ends by discussing these experiences and the current and upcoming challenges in Senegal.

2 Background

The work presented in this chapter is in relation to R&D projects initiated in Gaston Berger University in Senegal. These projects aim at providing solutions to the problems faced in data management and usage on the one hand, for the development of the Senegalese River Valley, and on the other hand, for the monitoring of disease spreading.

Fig. 1 Map of the Senegal river drainage basin (from wikipedia.org)

2.1 Information and Knowledge System for the Development of the Senegalese River Valley

As shown in Fig. 1, the Senegal River flows by delimiting a natural border between the Senegal, the Mauritania, and the Mali.

To make better use of resources available in the valley, the Senegal River's countries, namely Senegal, Mauritania, and Mali are organized within the Organization for the Development of the Senegal River (OMVS). So, under the direction of OMVS, two main dams, namely *Diama* and *Manantali*, are constructed and commissioned in respectively 1986 and 1988. These dams allow then to have total control over the river. Indeed, the available water reservoirs in the river must allow, for any season, an efficient irrigation of arranged areas. The climate around the river is Sahelian: rainfall is very low, erratic rains are distributed over both space and time,

temperatures are consistently high. Some soils on the edge of the river are salted and their culture requires important arrangements and expensive drainage systems.

The predominant crop in the Senegal River valley is rice. The other cultivated crops are mainly sugarcane and vegetable crops. The national SAED was established in 1964 in order to develop and supervise rice cultivation in the Senegal River valley. The main purpose is to reduce the country's dependence from imports. It is important to notice that rice is widely consumed by Senegalese population.

The development of the Senegal River Valley, the construction of the Diama (anti-salt dam near the mouth) and Manantali (hydroelectric dam and water reservoir), the development of irrigated crops, the intensification of agriculture, and increased water needs of urban populations have led to considerable development problems. So, the enforcement policy of the Senegal River valley seeks development and management models capable of achieving the ambitions of socio-economic development of the region. This cannot be accomplished without an efficient control of information and a rational exploitation of existing data on the region intended to be processed into relevant information and knowledge directly usable. Indeed, there is a significant available amount of data on the Senegal River valley. But these data are heterogeneous and scattered across multiple data sources and do not belong, mostly, to the same institution. Some data are even in a raw format and thus hardly exploitable.

Since several years, the development of the Senegal River involves experts from various agencies (ISRA—Senegalese Institute of Agricultural Research, SAED—National Society for Development and Exploitation of the Delta Lands, WHO—World Health Organization, FAO—Food and Agriculture Organization of the United Nations, etc.) working on many different skills (hydraulic, agricultural activities, mathematics, health, etc.). The work carried out by experts from these different organizations frequently result in the production and the exploitation of large amount of data. However, the management and the use of this critical amount of data are far from being satisfactory. The main reason is related to the nature of the data which are often voluminous, distributed, heterogeneous, and confidential. Furthermore, another major constraint is related to the increasing evolution of data due to hydro-climatic hazards, desertification, and permanent changes in the socio-economic layout of the region.

These problems should benefit from the latest achievements in the field of computer science research to have appropriate answers. In order to allow data producers and data consumers to deal with the most relevant knowledge from these data, it is important to provide them efficient and adapted means for collection, organization, management, diffusion, and exploitation of these data.

The Gaston Berger University of Saint-Louis in Senegal is playing a key role in developing knowledge management methods and tools adapted to the river enhancement problems. Thus, we initiated in 2005 the project SIC-Senegal (Information and Knowledge System). The main purpose of this project is to specify and implement an Information and Knowledge System for the development of the Senegal River. The SIC-Senegal platform must be accessible through the Web and allow mainly (1) the integration of information from heterogeneous sources and

(2) the use of integrated data to provide services and facilitate decision-making, as well as search engine for information retrieval from existing data. The project is conducted in partnership with some partners concerned by the development of the Senegal River. Such partnership already exists between the Numerical Analysis and Computer Science Laboratory of the Gaston Berger University and some of these organizations. Organizations such as the DRDR-SL (Regional Directorate of Rural Development of Saint-Louis) and WARDA (Rice Centre for Africa, Regional Station of the Sahel) agreed from the beginning to be our partners in this project.

The resulted information system can be used by all individuals and organizations involved in the development of the Senegal River: experts and researchers producing and manipulating data, as well as politicians and donors using these data for decision-making. The use of the system will facilitate the valorization of the region enabling solving development problems. This will have, therefore, important socio-economic benefits for the populations of the Senegal River region. SIC-Senegal also provides a tool which will allow teachers and researchers for developing models and validating research results.

It should be noted that such a system is easily adaptable and reusable by researchers and/or policy makers in other areas facing similar development problems.

2.2 Health System Management in Senegal

Great amounts of data are daily generated by medical activities such as medical consultation, hospitalization report, medical analysis, birth and death statement, etc. However, these data are essentially recorded in print medium and images whose access and use are mainly done manually. Nevertheless, some medical organizations have their own software for patient record management, appointment, wages, etc. without any possibility of sharing these data or communicating with other medical structures. This leads to some difficulties in reusing or sharing these data because of their structural heterogeneity.

Thus, we have proposed in 2011 to set up a Senegalese National Medical Information System (SIMENS) in order to overcome these problems. The purpose of SIMENS is to build a patient health record system. All the medical data of a specific patient can be accessed by medical practitioners at the national level by means of a unique identifier. All the data collected during medical activities will be recorded in a structured manner within the system. Then, they can be used for monitoring health status at the patient and population levels, risk analysis, early warning, disease spreading prevention and control, etc. The key to success this project relies on the patient personal data privacy management, the unification of data structure, and the rapidity of data integration at the national level for decision making.

The project also aims at recording all the public and private hospitals, specialist and general practitioners, medical laboratories, pharmacies, medical equipment within the medical organizations, etc. for information research and resources management purposes.

The general intended goal of the SIMENS project is to support the management and capitalization of medical information in Senegal. The project is driven through three main phases:

- Phase 1: development of a software for the daily medical activities management and medical data record.
- Phase 2: development of integration system of data produced at the hospital level by the software implemented in the phase 1.
- Phase 3: development of end-user tools for information research and decision making by medical and political authorities.

3 Agriculture Data Sharing: Ontology-Based Approach

The data Integration challenge that we address here is related to the enhancement of the Senegal River valley. We will particularly focus on the context of sharing agricultural data by experts and agencies involved in the development of this Senegal River Valley. Indeed, the entire region of the Senegal River valley is a large area with a great agricultural potential. It covers very fertile soils where are grown various crop in different agricultural seasons: irrigated, rainfed, flood-recession agricultures etc. Thousands of tons of crop varieties (rice, tomatoes, onions, sweet potatoes, sorghum, etc.) are thus produced in the valley each year.

To deal efficiently with the large amount of these heterogeneous agricultural data, many relevant works have been realized in the context of SIC-Senegal project [12, 21, 24, 29]. Thus, a first data integration approach has been proposed in [28]. In this work, data integration issue is addressed in a context of sharing agricultural data by several agencies involved as partners in the development of the Senegal River valley; each partner usually collects and stores their data in different formats (spreadsheets, relational databases, etc.). The proposed data integration process is then performed within each partner's source and is achieved in two phases. A first pre-integration phase allows the construction of an XML documents warehouse, called *dataweb*, for each partner. This phase relies on a previous work on XML-based data integration realized by [18, 22]. The authors define a dataweb as "an XML documents repository built from data resulting from heterogeneous sources." In the context of the SIC-Senegal project, dataweb approach is used to realize structural XML-based data integration for each partner. Each partner's dataweb is then materialized by XML data coming from its heterogeneous data sources. In the second step, a local ontology is built from the dataweb of each partner. Figure 2 gives an idea of how the concepts and relations that make up the local ontology are extracted from XML documents.

After this first integration phase allowing the building of local ontologies that semantically describes partner's data, [24] proposes a semi-automatic solution to resolve one of the important phases of the data integration process, i.e., the construction of the intended global ontology to be shared between all the partners. Furthermore, the works achieved by [24] also build a mediation system that provides necessary functionalities that enable partners to share data through an uniform query

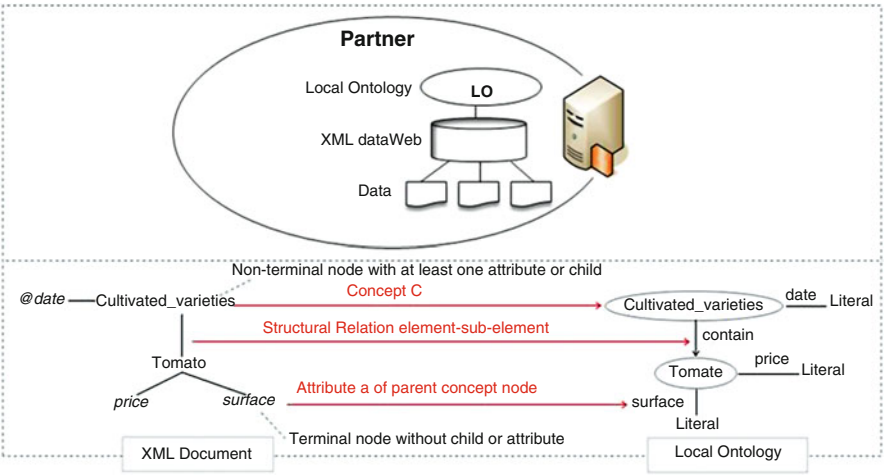

Fig. 2 Structural integration of a partner

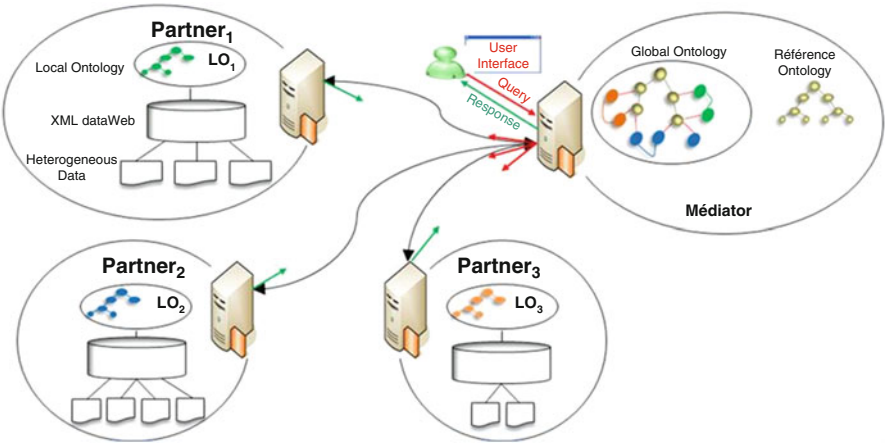

Fig. 3 General mediation architecture in the context of SIC-Senegal project

interface. The general architecture of this system that we call SIC-Agri (Information and Knowledge System for Agricultural data sharing) is shown in Fig. 3.

In this architecture, the global ontology is built semi-automatically from an existing domain-reference ontology. Our strategic use of this reference ontology in the building process allows us to make our mediator system (1) generic because it relies on a reference ontology representing the domain of the mediation, so it can be applied as soon as a reference ontology exists for the desired domain, (2) lightweight in the sense that the sources remain entirely autonomous from each other and the mediator, and (3) dynamic in the sense that updating, adding, and removing a source involves few computations.

4 Senegalese National Health System Project

Handling epidemiological phenomena needs to take into account all the disease dimensions and their risk factors. The dimensions depend on the scale on which the disease is observed and analyzed. Thus, we distinguish the three following dimensions:

- *The organic dimension* which is relevant to the biological perspective studying the interactions between the disease pathogen and the host organism.
- *The individual dimension* corresponding to the clinical perspective whose role focuses on patient care (diagnostic, treatment, nursing, etc.).
- *The population dimension* corresponding to the epidemiological perspective which aims at studying the disease spreading dynamics, prevention, and control strategies, etc.

Our actual projects are mainly focusing on the individual and population dimensions corresponding respectively to the clinical and epidemiological perspectives. In the next subjections, we are presenting:

- The SIMENS project which focuses mainly on the disease management at the individual level corresponding to the clinical perspective.
- The design of an ontology-based epidemiological monitoring system that was initiated in a PhD [4] and was applied to the schistosomiasis disease monitoring [6] in the Senegalese River Valley.

4.1 Senegalese National Medical Information System

In Senegal, information management within organizations and administrations has undergone lightning transformation through the development of Information and Communication Technologies (ICT). Areas such as larges retailers, banks, the private enterprises in general, have got robust information systems allowing them to take benefits from the data they produced. However, the public health area delays to take advantage from these innovative trends of ICT. But, great amount of data produced during the medical activities are still recorded in printed medium. These problems and limits have motivated the SIMENS project whose main objective is to build an application to assist the daily medical activities and to provide an electronic record system of the medical data. Beyond the management of the hospital daily tasks, SIMENS propose a medical data integration system from the hospital level to an upper level (departmental, regional, and national). Thus, it provides an overview and decision making tools to the medical and political authorities. Figure 4 shows the basic architecture of the SIMENS system.

The SIMENS project is currently supported by the Senegalese High Study and Research Ministry (MESR) through the Scientific and Technic Research Impetus Funding (FIRST) program. It is carried out by the Training and Research Faculty

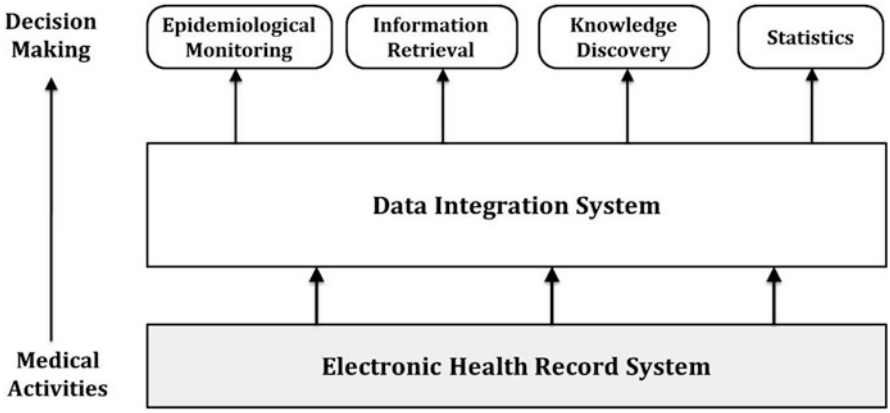

Fig. 4 Senegalese National Electronic Health Record System Architecture

(UFR) in Applied Sciences and Technology (SAT) and the UFR of Health Sciences (2S) at Gaston Berger University in close collaboration with the Regional Hospital Centre of Saint-Louis (CHRSL).

The electronic health record system comprises four modules:

- Care management system: consultation, diagnosis, hospitalization, medical appointment, medical analysis and exam, medicinal prescription, etc.
- Patient health record: patient health history at the hospital level for a specific pathology and at the national level at all time.
- Administrative and financial management: administrative patient record, medical personnel management, invoicing and wage, pharmacy, stock and supply, birth and death statements, etc.
- Web portal: patient and physician member area, medical structures research area (hospital, clinic, medical laboratory, etc.), national geographical information system (statistic, disease prevalence, etc.), etc.

Another important goal of the SIMENS project is the creation of a national patient record in a standardized format of the patient medical data. Thus, the patient can be identified in a unique manner in the whole country and his health history can be used for his treatment (intolerance and allergic alert during medicinal prescription) and also for his health monitoring (vaccine updates and appointment alert for instance).

4.2 Ontology-Based Epidemiological Monitoring System

Epidemiological monitoring systems are commonly developed to prevent and control disease spreading. This monitoring is carried out both to predict the

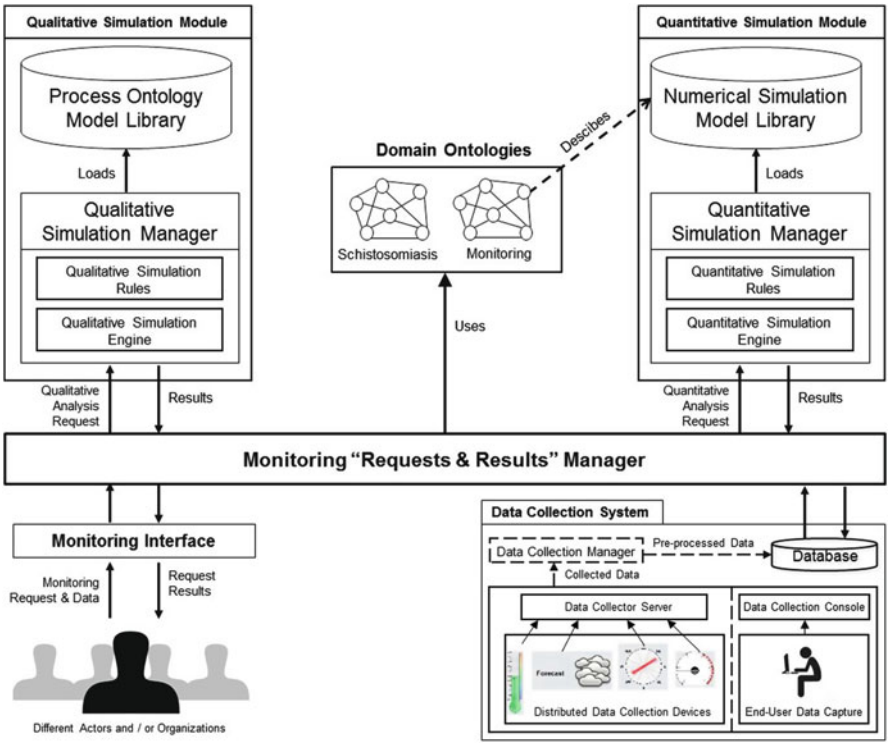

Fig. 5 Architecture of the schistosomiasis monitoring system in Senegal [6]

impact of events constituting disease spreading risk factors in order to control its progress and to suggest action plans to prevent identified risks [3, 5]. In the architecture that we have proposed (Fig. 5), risk prediction and decision making are performed using quantitative or qualitative simulations. These two types of analysis require the integration of data related to the disease, pathogen lifestyle, demographics, environmental, social and economic risk factors, etc. Figure 5 also shows the "Monitoring" module ontology, in the domain ontologies component, designed to support data integration and numerical simulation model composition during quantitative simulations for the schistosomiasis monitoring in the Senegal River Valley. The "Monitoring" module ontology describes the input and output data of all simulation models, the monitoring data sources, and the mapping and transformation rules applied before loading the data in the simulation platform.

Quantitative approaches are based on quantitative simulations using numerical models. These models built from epidemiological investigations allow explaining the dynamics of the spread of the disease and to validate assumptions on it. However, these models are difficult to use for predictions or decision making. Indeed, numerical simulation models require numerical data that are difficult to acquire in real time, especially in the developing countries. The qualitative simulation

approach is based on *process ontologies* and uses qualitative data as input. It allows for qualitative risk estimation when accurate and complete numerical data are not available. Qualitative simulations pull their data from messages posted in real-time in social networks such as twitter. The proposed ontologies in [4] are designed to take into account the description of the data and their sources to facilitate their integration in the simulation platform.

In this chapter, we focus on the challenges of the qualitative approach which represents an alternative and innovative solution because of the emergence of the mobile Internet in Africa and acculturation of users in social networks. Although the qualitative approach is considered as an alternative to the problem of quantitative data acquisition, it also presents some challenges. Indeed, it is necessary to determine the sources of these qualitative data and build a data retrieval and integration system from these sources. The data may be collected from designed persons as sentinels in the monitored areas to transmit data from their smartphones through an application to the data servers. They may also be retrieved from publications in social networks like twitter, reporting events in real time. Also, it is necessary to analyze text stream to determine the messages related to events that have been observed in the monitored areas and that constitute risk factors for a specific disease. Similarly, through facebook, beyond sharing of multimedia data such as images, we can notice that many users also report instantaneously events by adding potential useful tags. These means of qualitative data collection are promising in developing countries with the proliferation of smartphones and mobile internet access. However, textual analysis of these messages to extract the relevant qualitative data is the key of success to this new approach of qualitative risk estimation. There are many studies [19, 30] that would be interesting to explore. But the major challenge lies on taking into account the local languages used very often in text publication. In addition, these messages are also mixed with the official language which is French and foreign languages such as English.

Furthermore, we have shown the variety of quantitative or qualitative data needed for the analysis of risk of schistosomiasis spreading and decision making. However, these data are not very structured and easily accessible. A perspective of this work is to conduct a project of "Linked Open Data" in Senegal for two reasons: first, to promote standardization of data belonging to different organizations (health, demography, environment, etc.), and then promote access to prior linked data for more meaningful analyses.

5 Towards a Framework for Data and Application Sharing

Our ambition is to develop a platform for data and application sharing. We come back here on the methodological approach used for data integration within the SIC-Senegal project. We focus on the evolution of the framework and how applications design is taken into account in this new configuration of the platform.

5.1 The Dataweb Approach

The architectural choices for the implementation of the Information and Knowledge System (SIC) were guided by the specificity of the environmental information system. Environmental Information Systems (EIS) have some specificity [14]:

- Data come from a variety, heterogeneous, and distributed sources: databases, flat files, geographic information systems, etc.
- They must meet a very diverse audience (managers, policy makers, researchers, journalists).
- The data validity lifespan is relatively large.

Also, a conservation/observation type is generally used in the implementation of the EIS [11]. The conservatory aspect expresses the fact that there is acquisition and storage of data or documents and that this information is not regularly updated, but archived. This information is then returned to the end user. Observatory aspect expresses the fact that the users of the system may make observations (measurements, syntheses, research, extraction, etc.).

Taking into account the technological environment emphasized by the advent of the Web and the migration of information systems to the web, our proposal was based on a "data warehouse" approach [21]. It provides solutions to the problems arising from the migration of information systems to the Web (heterogeneous data integration, search for relevant information, diffusion of integrated data through the Web).

For the integration of heterogeneous data sources, we introduced the concept of *dataweb* which can be defined as a "XML documents warehouse" built from data of heterogeneous sources [22]. We proposed a dataweb model based on a global source. A global source incorporates, in a XML format, data from structured and semi-structured sources. It is completed by a catalog of metadata that enables the management of the dataweb (consultation, query, and update).

5.2 The Semantic Dataweb Approach

Using dataweb approach in the SIC-Senegal project has led us to integrate a semantic dimension in his definition. This allows us to solve structural and semantic heterogeneity of partner's data during their integration.

To solve the problem of semantic heterogeneity, our approach is to use controlled vocabulary describing the data of each partner as a support for the construction of an ontology called partner. Considering the environmental nature of the described data, this phase of semantic integration is mainly based on bottom-up and semi-automatic construction of ontologies describing the data of the sources through the reuse of existing ontology of the domain called Agricultural Ontology Service (AOS) [17].

To integrate data from partners, each dataweb is associated with an ontology. Thus, the concept of dataweb evolves into that of *semantic dataweb* [22, 28]. This is done by semi-automatic construction for each partner of (1) a specific OWL ontology from the XML data of his warehouse, (2) basis of annotations specifying (in RDF) relationship between his specific ontology and dataweb's XML documents, and (3) a generic ontology, containing the concepts of the partner's ontology that could be shared with others, in the context of mediation.

To perform the mediation of these semantic dataweb, we exploit the work described in [13]. It consists in using a hub-based system (or peer servers) whose architecture is the same for each partner. So, we give the end user the ability to query the metadata contained in the annotation database via SPARQL distributed queries using the global ontology. We propose a method to generate automatically a global ontology from local ontologies representing data sources to integrate. This is an automated construction, requiring extremely low human intervention, of a mediator system operating in an environment of semantic web type, which can be either an intranet or Internet. With its incremental construction and adaptation to the OBDA (Ontology-Based Data Access) principles, this mediator system has the advantages of being open, flexible, and scalable. In addition, this approach is generic in the sense that it is parameterized by a *reference ontology*, representing the area in which the mediation system is desired [25–27].

More generally, we consider, as it is the case in the SIC-Senegal project, the pooling of resources (data and applications) by partners located on remote sites. We are particularly interested in the dynamic discovery and invocation of shared services knowing that the necessary data for the performance of a service may be distributed among multiple partners or may be the result of the previous invocation of a service.

We are interested in problems related to data sharing (sustainability, accessibility, efficient transportation, security) and applications (discovery, invocation, composition, ergonomy). A particular focus is on optimizing the management of data needed for the invocation of a service especially when these data are in several different partners.

It is within this context that we are interested in [1], as many semantic web applications, to the issue of integration of data from distributed RDF sources. For the execution of distributed queries, for example, several solutions are proposed as SPARQL 1.1 Federation that defines a set of extensions to the SPARQL query language. If these extensions allow delegating part of the query to a set of services, the major problem is to determine automatically the services having the answers to a given query. Solving this problem is particularly crucial in a context like the Linking Open Data where many heterogeneous RDF sources are published. To send a query to only appropriate data sources, a complete and compact description of each source is required. To know the content of an RDF source we can rely on the graph patterns that compose it. In this work, we are interested in the extraction of these graphs patterns by proposing an extension of the DFS (Depth-First Search) coding system to the RDF graph patterns.

6 Related Work

Data Integration involves combining data residing in different sources and providing users with a unified view of these data. This process is important and appears with increasing frequency as the need to share existing data increases (in the commercial or biological domains for example). Indeed, many organizations hold some similar data in specific domains and want to share some parts of them (merging databases of similar companies or combining research results from different bioinformatics sources for instance). Data integration may then alleviate users from knowing the structure of different sources, as well as the way they are conciliated, when making queries [33], through the design of a global schema. In the context of Semantic Web, such a global schema is generally represented by an ontology. Indeed, the number of data sources describing their data with local ontologies is growing and the integrated access to heterogeneous data sources, called ontology-based data integration [20], is a challenging issue. Ontologies offer a formal semantics [15, 16, 31] which allows the automation of tasks such as heterogeneity resolution, consistency checking, inference, and global schema (ontology) construction. So, ontologies are used in many research works to solve data integration problems. In [2, 7, 8], ontologies have been used to resolve semantic heterogeneity problem by adding metadata to existing information sources, as in our approach. Other works are focused on environmental domain, particularly agriculture. In the Special Issue on Agricultural Metadata and Semantics,[1] some ontology-based solutions have been proposed for the management of agriculture knowledge and data. Ma [23] proposes an ontology-based approach for aligning whole farm plan data with other data collections of interest.

The originality of our approach in this proposal lies on the use of a reference ontology to automatically build a global ontology which is appropriate to the sources to be integrated as well as to the targeted area of application (represented by the reference ontology). The automation of building the global schema, as we did in [25, 27], is not much studied in the existing works.

7 Discussion

Notable advances have been made during the last decade in Senegal in the area of computing and ICT (networking of public institutions, increase in computer science graduated students, strong emergence of mobile telephony, etc.). These advances have led a lot of public and private organizations to implement information systems more or less operational. The implementation of these information systems should enable decision-makers having in real-time reliable data for designing and monitoring of public policies. However, these heterogeneous and distributed

[1] International Journal of Metadata, Semantics and Ontologies 2009 Vol. 4 No. 1/2.

operational information systems have led us to the need of data integration. Thus, we have focused our researches in proposing solutions for data integration and tools enabling knowledge management by giving priority to the areas of agriculture and health. For agriculture, we have proposed an ontology-based system for enabling actors involved in the development of the Senegal River Valley to share their information and knowledge. The health domain is concerned with the design of a national medical information system providing an integrated view of data related to medical activities and an ontology-based epidemiological monitoring system. These two applications developed within the context of R&D projects were initiated and carried out at the Gaston Berger University during the last decade. This shows our concern in providing through an ontology-based data integration approach a uniform access to distributed and heterogeneous data sources. In addition, it has been demonstrated that the building of the Diama and Manantali dams respectively in 1986 and in 1988 has caused the spread of schistosomiasis in the region and particularly the emergence of intestinal schistosomiasis in Richard Toll [32]. This demonstrates the consequence of not taking into account the interdependence between data from different but interdependent areas. This is one of the reasons justifying our goal to define an integrative data integration framework of data from different but potentially interrelated domains. This framework is based on an ontological layer whose objective is to facilitate the resolution of structural and semantic heterogeneity of data from different applications and databases.

Having noted the failure of several economic and social development models in developing countries, we believe that it is suitable to experiment new ICT-based approaches. We also believe that the key of success of these new endeavors lies on the promotion, preservation, and diffusion of the wealth of our socio-cultural background and the natural environment protection of our country. These are the reasons why we initiated the project "social and semantic web platform for sharing knowledge about communities" [9] supported by UNESCO and the AUF[2] funding programs. This research project fits with a need of a framework for sharing knowledge about our communities relying on social and semantic web technologies. These solutions must take into account local realities, especially regarding access to ICTs. For the same reasons, we are also involved actively in the CLUVA (Climate Change and Urban Vulnerability in Africa) project which was initiated in 2010 to enhance the understanding of climate change risks and vulnerability of African urban spaces. Knowledge management is a major challenge in this domain issues. Our main goals in this project were to bridge the gap of lack of knowledge management systems in this area by proposing ontology design patterns of the climate change domain [10] and a semantic-based architecture of multi-agent system.

Our perspective is to move towards a framework for sharing data and applications to enable, for example, the government and its partners or organizations and experts, to share data and applications in a distributed and secure environment.

[2] http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/pdf/SenegalGastonBergerFR.pdf.

Such a framework will be based on methods and tools for knowledge management (ontologies, semantic web standards, and technologies, etc.) and provides decision makers with the means to draw up public policies based on reliable and complete knowledge. Thus, our future work will be focused on providing a Linked Open Data Platform across the domain interest covered in this chapter in order to provide our country the means to face the need of public data integration and sharing.

References

1. A. Basse, Caractérisation incrémentale d'une base de triplets RDF (Thèse de doctorat unique). Université de Nice Sophia Antipolis et Université Gaston Berger de Saint-Louis du Sénégal (2012)
2. A. Buccella, A. Cechich, R. Cechich, N.R. Brisaboa, An ontology approach to data integration. J. Comput. Sci. Technol. **3**, 62–68 (2003)
3. F. Buton, De l'expertise scientifique à l'intelligence épidémiologique: l'activité de veille sanitaire. Genèses **4**, 71–91 (2006)
4. G. Camara, Conception d'un système de veille épidémiologique à base d'ontologies: application à la schistosomiase au Sénégal. (Doctorat Unique en Informatique). Universités Gaston Berger de Saint-Louis et Pierre et Marie Curie de Paris (2013)
5. G. Camara, S. Despres, R. Djedidi, M. Lo, Modélisation ontologique de processus dans le domaine de la veille épidémiologique, in: Actes de La Conférence Reconnaissance Des Formes et Intelligence Artificielle (RFIA2012) (2012)
6. G. Camara, S. Despres, R. Djedidi, M. Lo, Building a schistosomiasis process ontology for an epidemiological monitoring system, in *Innovations in Intelligent Machines-4*, ed. by C. Faucher, L.C. Jain. Studies in Computational Intelligence (Springer International Publishing Switzerland, 2014), pp. 75–99
7. I.F. Cruz, H. Xiao, F. Hsu, An ontology-based framework for XML semantic integration, in *Proceedings of the International Database Engineering and Applications Symposium*, *IDEAS'04*, 2004, pp. 217–226. doi:10.1109/IDEAS.2004.1319794
8. Z. Cui, D. Jones, P. O'Brien, Issues in ontology-based information integration, in *Proceedings of the IJCAI-01 Workshop on Ontologies and Information Sharing*, CEUR, Seattle, 2001, pp. 141–146
9. P.F. Diallo, S.M. Ndiaye, M. Lo, Study of sociocultural ontology, in *The First International Conference on Social Eco-Informatics (SOTICS 2011)*, Barcelona, 2011
10. I. Diop, M. Lo, An ontology design pattern of the multidisciplinary and complex field of climate change. Adv. Comput. Sci. Int. J. **2**, 104–113 (2013)
11. P. Dzeakou, P. Morand, C. Mullon, Méthodes et architectures des Systèmes d'Information sur l'environnement, in *Colloque Africain Sur La Recherche En Informatique et En Mathématiques Appliquées (CARI'1998)*, Dakar, 1998, pp. 509–520
12. D.C. Faye, G. Nachouki, P. Valduriez, SenPeer: un système pair-à-pair de médiation de données. ARIMA **4**, 24–52 (2006)
13. F. Gandon, M. Lo, C. Niang, Un modèle d'index pour la résolution distribuée de requêtes sur un nombre restreint de bases d'annotations RDF, in *Actes Des 19es Journées Francophones d'Ingénierie Des Connaissances (IC 2008)*, 2008, pp. 25–35
14. O. Gayte, J.-P. Cheylan, T. Libourel, S. Lardon, *Conception des systèmes d'information sur l'environnement* (Hermes, Paris, 1997)
15. T.R. Gruber, Toward principles for the design of ontologies used for knowledge sharing. Int. J. Hum. Comput. Stud. **43**, 907–928 (1995). doi:10.1006/ijhc.1995.1081
16. N. Guarino, Formal ontology and information systems, in *Proceedings of the 1st International Conference on Formal Ontologies in Information Systems (FOIS'98)*, ed. by N. Guarino (Ios Press, Trento, 1998), pp. 3–15

17. K. Hagedorn, F. Fisseha, Agricultural Ontology Service (AOS). A tool for Facilitating Access to Knowledge. UN FAO, Rome, 2002
18. A. Hocine, M. Lo, Modeling and information retrieval on XML-based dataweb, in *Advances in Information Systems*, ed. by T. Yakhno. Lecture Notes in Computer Science (Springer, Heidelberg, 2000), pp. 398–408
19. B. Klein, X. Laiseca, D. Casado-Mansilla, D. López-de-Ipiña, A.P. Nespral, Detection and extracting of emergency knowledge from Twitter streams, in *Ubiquitous Computing and Ambient Intelligence*, ed. by J. Bravo, D. López-de-Ipiña, F. Moya. Lecture Notes in Computer Science (Springer, Heidelberg, 2012), pp. 462–469
20. M. Lenzerini, Data integration: a theoretical perspective, in *Proceedings of the Twenty-First ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS'02* (ACM, New York, 2002), pp. 233–246. doi:10.1145/543613.543644
21. M. Lo, DataWeb Bases sur XML: Modélisation et recherche d'informations pertinentes. Université de Pau et des Pays de l'Adour (2002)
22. M. Lo, A. Hocine, ISYWEB: an XML-based architecture for web information systems, in *SITIS*, ed. by R. Chbeir, A. Dipanda, K. Yétongnon (Dicolor Press, Yaounde, 2005), pp. 116–121
23. H. Ma, Ontology-based agri-environmental planning for whole farm plans, in *Advances in Conceptual Modeling – Applications and Challenges*, ed. by J. Trujillo, G. Dobbie, H. Kangassalo, S. Hartmann, M. Kirchberg, M. Rossi, I. Reinhartz-Berger, E. Zimányi, F. Frasincar. Lecture Notes in Computer Science (Springer, Heidelberg, 2010), pp. 65–74
24. C. Niang, Vers plus d'automatisation dans la construction de systèmes médiateurs pour le web sémantique: une application des logiques de description. Université Gaston Berger de Saint-Louis, Sénégal, 2013
25. C. Niang, B. Bouchou, M. Lo, Y. Sam, Automatic building of an appropriate global ontology, in *Advances in Databases and Information Systems*, ed. by J. Eder, M. Bielikova, A.M. Tjoa. Lecture Notes in Computer Science (Springer, Heidelberg, 2011), pp. 429–443
26. C. Niang, B. Bouchou, M. Lo, Y. Sam, Querying a semi-automated data integration system, in *Database and Expert Systems Applications*, ed. by S.W. Liddle, K.-D. Schewe, A.M. Tjoa, X. Zhou. Lecture Notes in Computer Science (Springer, Heidelberg, 2012), pp. 305–313
27. C. Niang, B.B. Markhoff, Y. Sam, M. Lo, A semi-automatic approach for global-schema construction in data integration systems. Int. J. Adapt. Resilient Auton. Syst. **4**, 35–53 (2013). doi:10.4018/jaras.2013040102
28. O. Sall, Contribution à la modélisation de données multi-sources de type DATAWEB basé sur XML. Université du Littoral Côte d'Opale (2010)
29. O. Sall, M. Lo, F. Gandon, C. Niang, I. Diop, Using XML data integration and ontology reuse to share agricultural data. Int. J. Metadata Semant. Ontol. **4**, 93–105 (2009). doi:10.1504/IJMSO.2009.026258
30. R. Steinberger, M. Ehrmann, J. Pajzs, M. Ebrahim, J. Steinberger, M. Turchi, Multilingual media monitoring and text analysis – challenges for highly inflected languages, in *Text, Speech, and Dialogue*, ed. by I. Habernal, V. Matoušek. Lecture Notes in Computer Science (Springer, Heidelberg, 2013), pp. 22–33
31. R. Studer, V. Richard Benjamins, D. Fensel, Knowledge engineering: principles and methods. IEEE Trans. Data Knowl. Eng. **25**, 161–197 (1998)
32. I. Talla, A. Kongs, P. Verlé, Preliminary study of the prevalence of human schistosomiasis in Richard-Toll (the Senegal river basin). Trans. R. Soc. Trop. Med. Hyg. **86**, 182–191 (1992)
33. H. Wache, T. Voegele, T. Visser, H. Stuckenschmidt, H. Schuster, G. Neumann, S. Huebner, Ontology-based integration of information: a survey of existing approaches, in Proceedings of *IJCAI-01 Workshop on Ontologies and Information Sharing*, Seattle, WA, USA, ed. by H. Stuckenschmidt (2001), pp. 108–117

Information and Communication Technology: A Tool for Health Care Delivery in Nigeria

Peter Adebayo Idowu

Abstract Information and Communication Technology (ICT) is arguably the most rapidly growing segment of the world ecosystem. ICT has become a major tool in delivery of health services and has had a revolutionary impact on how we live and perceive the world. It has given birth to the contemporary "Es" such as e-learning, e-commerce, e-governance, e-banking, and e-health. This chapter discusses the state of health informatics in Nigeria with some of the existing software developed for health care delivery in Nigeria. It also discusses the benefits that can be derived from the use of ICT in health care delivery in Nigeria. The chapter finally discusses the problems associated with the use of ICT in health care delivery in Nigeria.

1 Introduction

Information and Communication Technology (ICT) is arguably the most rapidly growing segment of the world ecosystem. The development in the sector pervades every human activity: social, economic, cultural, religious, political, or health care [20]. The huge networking possibilities afforded by ICT have significantly transformed the health care systems in the world [11, 25] dispersing health care information with comparative ease, bringing patients closer to care givers, making access to the best health care technology, and expertise available to the remotest parts of the world.

ICT enables people to interact and communicate no matter the distance. The world in which we live in today has been changed by ICT. It has the potential to transform radically the delivery of health care and to assist in defining strategies to address future health problems. However, the rate of deployment of the technologies has been lower in developing countries, especially those of sub-Saharan Africa when compared with developed world. ICT has assisted in driving down healthcare

P.A. Idowu, Ph.D. (✉)
Department of Computer Science & Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria
e-mail: paidowu@oauife.edu.ng; paidowu1@yahoo.com

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_4

costs [31] and has improved the delivery and effectiveness of healthcare services through help in disease management, improved patient safety, and decision support for practitioners [28].

In Nigeria, most government administrations have promised to address the problem facing the delivery of health care in the country for the past decades now with the attendant opportunities, challenges, and achievements. Yet, it seems there is no significant change in the delivery of health care and this is not far from the fact that the government of Nigeria has failed to fully utilize the opportunity offered by ICT in health care sector. Consider the use of life expectancy as a measure of the state of the effectiveness of the health sector of a nation. According to World Health Organization [35] report that was released in the first half of 2009, the life expectancy in Nigeria was noted to have moved from 47 to 49 years [35]. When compared to other countries where the life expectancy goes as long as 83 years, and how well they are doing based on life expectancy measures, it can be seen that a lot more remains to be desired in the state of health care and services delivery in Nigeria.

Despite Nigeria's claims to have reorganized its health care delivery system since Bamako's 1987 initiative, which focused on how to increase access to good health care systems, the government is yet to implement any meaningful ICT policy in the health sector. It is obvious that the regions of the world where the life expectancy of nations are higher also are the ones where the use of ICTs are becoming more effectively mainstreamed in the different areas of national life including the health sector. It is obvious that there is a significant correlation between the level of the use of ICT in nations and the level of performance in their health sectors. Going by the immense contributions of ICT in developed nations, Nigeria must consider a more effective deployment of ICT in its health delivery system [29]. So there is a serious need for the nation to fully deploy ICT in health sector and quickly tackle some of the challenges facing the proper deployment of ICT in the sector and the nation as a whole.

The rationale for this chapter is to discuss the state of health informatics in Nigeria with some of the existing software developed for health care delivery in Nigeria. In this chapter benefits that can be derived from the use of ICT in health care delivery in Nigeria are also discussed. The chapter finally discusses the future trends of ICT in health care delivery and the problems associated with the use of ICT in health care delivery in Nigeria, and it offers possible suggestions on how it can be addressed.

2 ICT in Nigeria

ICT is a tool for capturing, processing, analyzing, storing, retrieving, transmitting, and presenting data and information for external uses with little or no human intervention. ICT is the integration of computing and telecommunication in order to process data and information for useful decision making.

The advent of ICT in many developing countries can be linked to international organizations (such as World Health Organisation, World Bank, United Nations Development Program, among others), research institutes, and educational institutions. For example, in 1993 Mongolia was provided with personal computers funded by the World Health Organisation to support health care services in the country [8]. In Nigeria, the history of ICT initiatives can be traced to 1950s when electronic and print media were formed, but there was little progress until National Broadcasting Commission (NBC) and the Nigerian Communication Commission (NCC) were formed in the early 1990s. In the late 1990s, NCC licensed some wireless telephone operators and the commission was reconstituted in 2000s so as to function well [10]. A year later after proper commissioning of NCC, the Obasanjo administration established the National Information Technology Development Agency (NITDA) to see to implementation of ICT policy in Nigeria, although this has little significant effect on the Nigerian economy.

In many industrialized countries of the world, there is a huge investment of resources into ICT in health care as a commitment to providing the most efficient and effective health care services to their teeming population. This is yet to be seen in many countries of sub-Saharan Africa. In Nigeria, knowledge and utilization of ICT is still very poor [7]. Access to the three components of ICT (computers, Internet, and telephones) is still largely limited to the elites and corporate institutions [20]. Telephony is the most developed and most accessible component of ICT among the general population in Nigeria. Even then, many public institutions including hospitals still lack effective internal and external telephone interconnectivity within and with the outside community respectively. Since the year 2001, there has been a revolutionary growth in the telephony industry in Nigeria [15, 16, 20]. This is expected to have permeated the health care institutions with significant positive impact on health care delivery.

Two years after the introduction of GSM in Nigeria, a study [15] was conducted among medical practitioners in some Nigerian teaching hospitals to find out the current level of utilization of telephones for health care delivery. The study observed that medical experts use their personal mobile phones to facilitate patient care at their own expense. It also revealed that telephony is the most useful ICT tools among medical experts in Nigeria as a means of facilitating health care delivery. It was reported that telephone is used in contacting wards and laboratories for vital information, materials, and equipment and calling on medical colleagues for second opinions. The study revealed that individuals' personal mobile phones were being used without any form of reimbursement from hospital management.

Both the government and hospital management had failed to take up the responsibility. Currently, the number of GSM service providers in the country has increased significantly and all of them keep rolling out several service options or packages including Internet services with varying benefits to attract customers. The extent to which the growth in telecommunication has impacted tertiary health care and how far this has enhanced the development of an ICT-driven health care practice in Nigeria is still to be known.

The Nigeria Federal Ministry of Health (FMOH) set up its ICT committee in 2003 with one of the goals being to achieve electronic linkage of the ministry with its hospitals, and the aim was to create a Wide Local Area Network (WLAN) whereby the tertiary hospitals can all access each other, exchange information, and create a reliable central database [5] but up till now that is yet to be achieved.

3 Health Informatics in Nigeria

Nigeria has one of the highest mortality rates in sub-Saharan Africa [34] with a population of 140,003,542 which is the highest in Africa but with inadequate health care delivery system. According to Hyman & Silver [13], inadequate health care delivery might give birth to health risk which means that Nigerians may be liable to health risk [13]. The problem of inadequate health care delivery is arguably minimized in countries like United Kingdom (UK) whose health services make more extensive use of ICT.

Health informatics which was formerly known as medical informatics was defined by Warner et al. [33] as the science that deals with the use of computers and communication technology to acquire, store, analyze, communicate, and display medical/health information and knowledge to facilitate understanding and improve the accuracy, timeliness, and reliability of decision making [33]. According to O'Carroll et al. [28], health informatics was defined as a demonstration of how organizations can use IT to bring their strategic goals from theory into practice [28]. The term medical informatics was changed to health informatics or healthcare informatics when some health officers felt that the term medical informatics had to do with physicians alone although today some people still use the two terms interchangeably. In this chapter, health informatics is defined as the use of telecommunication and computer to capture, process, analyses, store, retrieve, disseminate, and present health-related information for the purpose of making useful health-related decision.

The history of health informatics in Nigeria started in the late 1980s when a collaborative research project between the Computing Centre of the University of Kuopio, Finland and Obafemi Awolowo University, and Obafemi Awolowo University Teaching Hospital (OAUTHC), Nigeria [15] was initiated and this initiative was part of INDEHELA (a long-term research project on Informatics Development for Health in Africa) details of which can be found at http://www.uku.fi/indehela/. The joint project produced a very rudimentary hospital information system based on the Veterans Administration's (VA) Admission Discharge Transfer, running on a stand-alone PC, which was in use at OAUTHC in 1991 [9]. According to Daini et al. [9], the group then organized the first International Working conference on Health Informatics in Africa which was held April 19–23, 1993 at Ile-Ife, Nigeria [12]. In the late 1990s, the Finnish/Nigerian research team decided to expand their rudimentary hospital information system with the aim of developing a comprehensive system suitable for use in all Nigerian teaching hospitals and

medical centers. The plan then was that by 2001, all the teaching hospitals in Nigeria would have Health Informatics units which could make use of standardized software which is yet to be realistic in Nigeria.

In reality, though some hospitals have computer or IT units, these serve primarily to support word processing for typing pools and offices. Development of the commercial software "Made in Nigeria Primary Healthcare and Hospital Information System" (MINPHIS) ran alongside the doctoral programme of one of the staff of the Department of Computer Science & Engineering, Obafemi Awolowo University and was completed in 2004. Unfortunately, the system was not tested at OAUTHC and only five teaching hospitals and medical centers use the system as at 2007. The primary reason for this limited use is the cost of purchasing the commercial software. As Finnish/Nigerian research team were working on primary healthcare and hospital information system, a Norwegian and South African team focused on a district-level information management system [3].

Since the advent of ICT in health sector in Nigeria, various software have been developed to aid good health care delivery and some of the computer based systems are briefly discussed below.

4 Related Works

In the literature, some authors have identified ICT as a tool for proper healthcare delivery, though most of the developing nations are facing some challenges with the proper implementations of ICT in health sector. According to Omotosho et al. [30], ICT has contributed immensely to improvement in health sector in different parts of the world and it reduces cost of healthcare delivery. The paper presents a framework for implementation of ICT in developing countries with evidences to why developing nations must embrace the opportunity [30].

In Ethiopia, about 85 % of people in the country live in remote areas and the country has shortage of medical experts. The available medical experts prefer to stay in urban areas. According to Shiferaw and Zolfo [32], these challenges give birth to the need for developing telemedicine in the country. However, there are problems in setting up and implementing a telemedicine program in the country [32].

In Nigeria, Idowu et al. [17, 18] conducted a survey to know the degree and the extent of incorporation of Information Technology in the Nigerian health sector and derived IT infusion models for popular IT indicators that are in use in Nigeria (Personal computers, Mobile phones, and the Internet) and subsequently investigates their impacts on the health care delivery system in Nigerian teaching hospitals [17]. The results of the study revealed that out of the three IT indicators considered, mobile phones are fastest in spreading. It also revealed that computers and mobile phones are in use in all the teaching hospitals.

In this chapter, benefits that can be derived from the use of ICT in health care delivery and the problems associated with its use in Nigeria are clearly discussed.

5 Existing Software Developed for Nigerian Health Sectors

There are quite a number of existing systems developed for health care delivery in Nigeria by different researchers in the country, but some of these systems are not embraced in Nigerian hospitals. The systems are briefly discussed below.

5.1 Made in Nigeria Primary Healthcare and Hospital Information System

MINPHIS is a product of research collaboration between the Computing Centre of the University of Kuopio, Finland and Obafemi Awolowo University, and OAUTHC, Nigeria.

The MINPHIS application keeps patient records and generates various reports for health management and research purposes. The reports include the patient status, medical history, and admissions plus indicators like length of stay per patient, discharge summaries, mortality and morbidity data, and operations. The application can answer ad hoc queries from medical researchers (e.g., cases of cholera for a period per geographical location for specific age group or sex or both). It can also provide performance information relevant to particular health care professionals, such as the mortality rates for patients treated by a particular staff member [2].

5.2 State Hospitals Network

State Hospitals Network (SHONET) is a distributed resource sharing system for hospitals in Nigeria. It employed the state hospitals in Osun State Nigeria as a test-bed. The system allows specialized hospital equipment and medical personnel to be shared by hospitals within the network. Network software comprising two major modules was developed using Linux and Windows 2000 Professional as platforms. The two modules are the File menu and the Doctor's Reference modules. The design principles include simplicity, flexibility, modularity, and expandability. The system has the potentials of increasing the productivity of medical personnel and enhancing the quality of medical diagnostics tremendously. The system was developed so that specialized hospital's equipment and medical officers can be shared over the network. Specialized equipment will be in different location based on diseases or health challenge peculiar to each location.

5.3 Hospital Referral System

Hospital Referral System (HORS) is an electronic referral system for hospitals in Nigeria. It is complementary to the SHONET facility. HORS allows an electronic-based medical consultation and referral system. The details of the hospital records or case notes and records of medical tests can be transmitted electronically from one hospital to another with better facilities, competence, and experience. The remote consultant therefore can examine the patient's case file and make appropriate prescriptions which are also transmitted electronically over the network to the originating hospital. This has helped in reducing traveling time for critical patients and also reduces death rates occurring as a result of lack of prompt attention. The prototyping of the system was successfully implemented [16].

5.4 Radiological Information System for Primary Health Centres

Radiological Information System for Primary Health Centres (RISP) focuses on radiological information. In order to increase the access and quality of medical care in remote parts of Nigeria, RISP was developed using Microsoft Visual Basic version 6, an object oriented programming language. Radiological images were transmitted over the network and the remote doctor got consultation and interpretation on the images sent to consultant radiologist in the tertiary hospital. RISP allows doctors in rural area to send radiological images to consultants in the teaching hospital for interpretation and/or consultation. It also reduces time, cost of traveling, and allows better management of patients. It also increases exposure of expertise and reduced anxiety where second opinion is rapidly provided and on occasions, surgery or other procedures are avoided for rural medical personnel [18].

5.5 Wireless Data Processing Model in Hospital Environment

This system is a wireless data processing model for hospital environment using Obafemi Awolowo University Teaching Hospital as case study. It makes use of Client Server architecture. Base station serves as the server while the mobiles serve as clients. As Medical personnel move from ward to ward, i.e., from one cell coverage to another, the drop in signal is not due to momentary fading but as the terminal is moving away from the base station. Frequency Hoping spread spectrum is employed for spread spectrum modulation. The channel access method employed is Fixed Assignment Access Method and the channel management employed is Code Division Multiple Access to allow multi-user operation. The data in the network is secured through encryptions and cryptography [19].

5.6 GSM Based Referral System for Primary Health Care Centers in Nigeria

This system focuses on use of GSM for hospital referral system in Nigeria. In an attempt to reduce paper work, eliminate the problem of case notes getting lost in transit, shield the contents of case notes away from patients which could also lead to a psychological breakdown by the patient, and as a result of all these reduce mortality rate in the primary health care centers in Nigeria, an electronic referral system is designed to alleviate all these aforementioned problems. Medical personnel (such as nurses, local health officers, and doctors) were interviewed in an attempt to acquire a detailed knowledge of their manual process of referring patient. The results showed that patients' case notes (which encompass patients' symptom, diagnosis, medication with the clinical number) were transmitted using mobile phone on a Global System for Mobile Communication carrier from the referral package within few seconds. In conclusion, the work has the potential to increase medical personnel productivity, reduce prenatal and neonatal mortality rates, improve medical care, and minimize the cost of referral since GSM facilities are already on ground and some mobile operators offer free text services [14].

5.7 Spatial Predictive Model for Malaria in Nigeria

This system is defined with an inbuilt malaria predictive model. Malaria is one of the diseases that contributed to health problems in Nigeria and this research developed a Geographic Information System (GIS) forecasting model for prediction of malaria in Nigeria. This research develops a GIS based prediction model for malaria in Nigeria. With the application of GIS in malaria control, Nigeria can move towards certain reduction (if not complete eradication) of malaria infection and deaths through the use of analysis and prediction. This system will enhance the efficacy of prevention efforts and will substantially reduce costs of prevention and treating malaria in Nigeria. The system will also serve as a good decision support system for public health officers, government office, and decision makers [21].

5.8 Web-Based Environmental Health Tracking System for Nigeria

The research developed a web-based environmental health tracking system for Nigeria. It addresses the problem of environmental health monitoring system facing Nigeria as a whole. Environment and the factors that are associated with it are the root causes of many epidemic diseases both in the developed and developing nations. In Nigeria, environmental health problems arise from population pressure

on housing, poor environmental sanitation, coupled with lack of safe drinkable water and basic housing facilities. Despite the deplorable state of environmental health (lack of clean and safe drinking water, bad housing condition, and so on), there is no reliable and timely means of surveillance or any monitoring system. The result of the work makes it possible for environmental health workers to capture environmental health situation of any house in Nigeria real time while on the field [22].

In addition, the system can handle waste management and solid refuse from their sources of generation through storage, collection, transportation, recovery, and treatment processes to disposal. With the system, an effective waste management system and a spatial view of waste collection locations in any local government area in Nigeria can be done. The use of this system will ease the job of the waste management unit of the local government areas in Nigeria in achieving a clean environment and mitigate the spread of epidemic in a way to ensure safety of all and sundry [23].

5.9 Summary of the Existing Software Developed for Nigerian Health Sector

SN	Scope	Strengths	Limitations	Remarks
i.	Keeps patient records and generates various reports for health management and research purposes	Handle the patient status, medical history, and admissions plus indicators like length of stay per patient, discharge summaries, mortality and morbidity data, and operations	Not web application. Not support inter hospitals	It will be better if it can link hospitals
ii.	Distributed resource sharing system for hospitals	Allows specialized hospital equipment and medical personnel to be shared by hospitals within the network	Cannot capture information about any diseases	It will reduce cost of running hospitals but need to handle other things in the hospital
iii.	Electronic referral system for hospitals	Case notes and records of medical tests can be transmitted electronically from one hospital to another hospital	Tracking of any diseases is not possible	It will improve mortality rate but need to handle other things in the hospital

(continued)

(continued)

iv.	Remote consultation by radiologist	Allows doctors in rural area to send radiological images to consultants for interpretation and or consultation	Remote consultation of radiological images only	It can link radiology clinics especially to handle breast cancers
v.	Wireless access by medical officer in the hospital	Allow multi-user operation in the hospitals	Limited to a single hospital	Mobile phone operators have provided better network service
vi.	Hospital referral system based on GSM	An electronic referral system to aid health care in rural areas	Based on GSM designed for rural dwellers	Remote areas can still make use of mobile network Internet facility
vii.	System to predict likelihood of malaria	Support spatial query	Is limited to malaria only	Is for the use of disease surveillance officer
viii.	Tracking of environmental health-related diseases	Support spatial query. Handles waste management	Limited to environmental health-related diseases and waste management	Is for the use of environmental health officer

Despite the fact that there are some hospitals or health-related applications developed for the use of Nigerian health sector, almost all of these applications are not in use in the country. Though the applications are developed to handle different areas of the health sector, they are not in use.

Also, the applications could be merged together to become one which will eventually handle most of the needs of the health sector in the country.

6 Benefits that ICT Can Offer in Health Care Delivery in Nigeria

ICT has benefited the health sector both in developed nations and developing nations, and the benefits affect the hospital's stakeholders. In this chapter, the hospital's stakeholders are grouped into three namely; patients, hospital managers, and health personnel. Patients are people receiving treatment or register to receive medical treatment in the hospital. They are generally recipient of health care services. Hospital managers are also known as healthcare administrators; they are the leaders, administrators of hospitals. The health personnel are the health workers and they comprise medical doctor, nurses, medical record officers, and medical laboratory scientists among others.

6.1 Patients

With the introduction of ICT in hospitals, patients can benefit in the following ways:

Appointment Booking

In 1993, the Federal Government of Nigeria directed the Federal Ministry of Health to start the NHIS [26] in the country in 1999 and the scheme was modified to cover more people [1]. NHIS became operational after it was officially launched by the Federal Government in 2005 [24].

The purpose of NHIS is to ensure the provision of health insurance "which shall entitle insured persons and their dependents the benefit of prescribed good quality and cost effective health services" [26]. Because of the cost effectiveness of the scheme, there are more beneficiaries that are entitling to the medical benefit. With this scheme, every Nigerian can choose any hospital to go for medical treatment, and this causes some problems in the Nigerian teaching hospitals, because there are many outpatients that are on queue at the clinic to compete for the limited appointment opportunities with physicians [4]. As a result of this, there is long waiting time before an NHIS outpatient can have access to Doctor or Physician.

Thus access and appointment availability become the key component of outpatient satisfaction within the hospital. With the electronic-based appointment booking system, the patient can book appointment with the doctor right at home or office and select the time that is convenient to see doctor.

This system will eliminate long waiting queue which is one of the major problems facing many outpatients in the Nigerian government hospitals [27]. It will also reduce the number of missed appointment and unnecessary outpatient queue at the clinics. The e-booking system will enable the patient to login with their hospital number to appointment booking interface and view the scheduling calendar for the available physician, available time, and room number of the available physician.

So, if Nigerian government can start implementation of this type of system in government hospitals, it will really assist the patients to focus more on their businesses and jobs which would result in more productivity, both for the patients and the employers of the patients.

Emergency Ambulance

In most developed nations, there are emergency ambulances which transport patients from their houses and accident locations to the nearest hospitals. These emergency ambulances are reached through a specialized number which is free from either mobile phone or land lines. With the use of this facility, a lot of lives have been saved. If Nigeria government can make use of this method too, it will really help in effective health care delivery system. The use of mobile phone is common

among most Nigerians; more than 70 % Nigerians have mobile phones. If Nigerian government can purchase ambulance and station them in some strategic areas, it will reduce death rate as a result of road accidents.

In South Western Nigeria, most states in the region have emergency ambulances stationed in some strategic places for transportation of accident victims to nearby hospitals for proper attention. Immediately after an accident occurs, the other road users or people passing by or any of the accident victims may just call the phone number of the emergency ambulance and on arrival it would transport the accident victims to nearby hospitals.

So, if Nigerian government can imitate these selected states and have emergency ambulances which can be contacted through mobile phones on all the federal roads and important locations, it would assist in reducing death rate caused by road accidents and also increase proper health care delivery because any hospitals that the government emergency ambulances get to (most especially government hospitals) the medical officers on duty would attend to them promptly.

Electronic Referral

Electronic referral which is also known as e-referral is an electronic platform that enables the seamless transfer of patient's information from one hospital or clinic to another. Electronic referral at the patient level ensures significant improvements to follow up care coordination by the creation of accurate and timely referrals.

In the UK, by the end of 2014 a new NHS e-referral service will be launched to replace the current choose and book service. The development of the e-referral system was due to the feedback from the patients and NHS professionals. The new referral system which will use enhanced technology will improve the quality of the referral experience for the patient, clinicians, and administrative staff.

The NHS e-referral system will help create more patient centered, people-powered services, thereby making hospital businesses much easier. In most hospitals especially in developing nations like Nigeria, referrals of patients are done through patient's referral letter which is paper-based. Most of the time, due to tight schedules, the doctors may not have time to write referral letter which may cause death of a lot of patients. The patients who leave one hospital to another without referral letters may eventually die because the new doctor would not have adequate information about the drug previously given to the patient. The new doctor may likely repeat the same drug, and which may lead to overdose in the patient and a lot have lost their lives as a result of such. So if there is electronic referral system which links all the hospitals in the country, especially all the NHIS hospitals, the new doctor would be able to identify the drug the patient has taken and know how to address the problem if the sickness persists.

If Nigerian government uses the opportunity offered by ICT in providing e-referral system, referral of patients from one hospital to a specialized hospital

would be easier and efficient. Communication of clinical data would be secured, referral letters would not be lost, and this would improve patient safety and quality of care given among other things.

Electronic Patient Case File

With the use of ICT in the hospitals, the patients' case files will be created, populated, modified, and deleted electronically. The result of patients' tests can be added to the patients' case files as soon as they are ready. As a result of this, patients' case files will not be lost even if patients do not come to hospitals for years, which is not the case with paper files. If a patient fails to come to the hospital for certain period of time, the file may be misplaced or lost.

So, if Nigerian government can make sure all the patients' case notes are in electronic form, it will really help the government to have proper health record of all people living in Nigeria. In all, it will make it easy for the government to be able to have a national database for diseases or health challenges peculiar to all the people living in Nigeria at any particular period of time.

6.2 Medical Personnel

In this chapter, the medical personnel include Physicians, Nurses, Pharmacist, and Medical Record Officers.

Electronic Referral System

It aids medical decisions as physicians have full patient information and share it from one physician to another. It is faster, secured, and remove delays related to paper based referral system. It also improves practice productivity. It improves quality of documentations by removing the use of illegible handwriting. On the long run, e-referral is cheaper than paper based referral. With the use of electronic referral system, physicians would be having access to patients' case file electronically in real-time. The physicians and researchers can query health problems or diseases that are peculiar to each community, which will eventually aid in decision making.

Research

Due to the complexity of human body and the continual emergence of new diseases, research in medical sciences has been an ongoing activity and ICT has changed the way medical scientists conduct researches. With ICT, a wide variety of medical information is online such as; symptom diagnosis, (for people who wish to identify

their problems without consulting a physician) access to medical research/journals, information about side effects, and information about available treatments among others. So, with ICT, one can access real-time information, get support from sufferers of similar ailment, and get access to relevant research results from other professionals, which can help cure an ailment.

In addition, since the use of ICT allows electronic keeping of patients file, it makes research easy especially in sub-Saharan Africa region, where there is almost no proper and efficient database of any disease. With the use of ICT, information about diseases at any particular time would be available and it will aid researchers in getting data to conduct researches. For example in Nigeria, it is almost difficult to get accurate data about malaria which is the commonest diseases in the country.

Despite the fact that Nigeria is among the first five nations of the world with highest number of people living with HIV/AIDS, but there is no existing national database, so to be able to know the spatial distribution of this deadly disease, but with the use of ICT in Nigerian hospitals, all these problems would be averted.

Medical Equipment

Nigerian doctors often complain of lack or inadequate equipment to work with. Whatever equipment they have is either outdated or antiquated. Today, equipment such as body scanner is now ICT-aided and can be used to process data into readable formats. Microprocessors can now be used to control a variety of medical devices such as pacemakers, and computer-guided lasers are used in very delicate operations such as removal of brain tumors.

The advantages are numerous—such computer-aided equipment save a lot of time and reduce waiting time. Interestingly, surgeons can now perform operations remotely through teleconferencing and prescribe drugs through telemedicine.

6.3 Hospital Management

The use of ICT in health sector reduces the cost of running hospitals [31]. For example, The NHS in UK has its own network known as NHSnet which has benefited all parts of NHS. Apart from data networking and Internet, the NHS spends millions of pounds each year on telephone services across England to aid patient transport services and emergency ambulance because the two areas help the NHS to deliver good health care facilities [27]. ICT introduces potential of sharing of patients files easily without any threat to patient privacy. It is used for hospital management such as admission and appointment management. It would also help the management of the hospitals to be able to identify what is needed at every period of time. The management would be able to know when to order to a particular drug and the quality of the drug to order for at any particular period of time.

6.4 Government

Gross Domestic Product

Most Nigerians who seek medical treatment overseas are merely enriching the purses of those countries and this has negative effect on Nigeria economy and of course, has negative effect on the country's Gross Domestic Product (GDP). So with the use of ICT in Nigerian health sector, more job opportunities would be created among the youths and such Nigerians who travel abroad would otherwise spend the money in Nigerian hospitals and this would increase the nation's GDP.

Employment

Unemployment remains one of the most critical problems facing the country today. Nigeria is the most populous country in Africa and second largest economy in the continent. Yet 80 % of Nigerian youths are unemployed [6]. Presently, most of the security challenges facing the country were attributed to the rate of unemployment among the youths in the country. The use of ICT in health care delivery system would create more jobs for millions of Nigerian youths that are jobless. Indeed, ICT will create a new vibrant economy in Nigeria which will generate good paying jobs for IT professionals, software developers and researchers, health economists, and insurers, who will significantly contribute to our GDP, expand our middle class, and help reduce our dependence on oil.

Aids in Decision Making

With the use of ICT in health sector, national disease database would be readily available and spatial distribution of each of the diseases can be viewed by the stakeholders. This would help the government to know diseases that are peculiar to each geopolitical zone and also aid the government in decision making. With the use of ICT health-related decision making would be easy because, all necessary data about patients and hospitals and facilities that are available in the hospitals and the required hospitals would be available in real-time online.

This would assist the policy makers to know the location to provide with what and what at a particular time. This would eventually lead to improvement in quality of health care services in the country. It would allow policy makers in government the ability to gather, collate, and analyze data retrospectively and in real time to allow for efficient budgetary allocation of scarce funds.

7 Challenges Facing the Use of ICT in Health Care Delivery System in Nigeria

Some of the nations within sub-Saharan Africa region are facing problems preventing them from fully utilizing ICT in health care delivery. Idowu et al. [20] identified some problems facing the use of ICT in Nigeria health care delivery [20]:

7.1 "Epileptic" Electric Power Supply

Any country that cannot provide uninterrupted Power Supply (UPS) to its citizens will definitely have problems with deployment of good ICT services. Nigeria is facing this problem. People brought up in developed nations of the world will find it difficult to adapt to the "epileptic" electric power supply in Nigeria, which has caused a lot of damage to research institutes computer laboratory equipments, as computer components such as hard disks and motherboards can be destroyed by interrupted and unreliable power supply. ICT equipment is made to function with other infrastructure such as electricity under "controlled conditions," i.e., stable and constant electricity supply. Most Internet facilities in Nigeria suffer frequent downtime due to power interruptions and equipment damage.

A few years ago, Ghana celebrated 1 year without power failure; in contrast, Nigeria, despite its claims to be "'the giant of Africa" rarely has stable and reliable electricity for ten consecutive hours. In developed nations, many companies supply and guarantee electricity and there is no problem of electric power supply.

Nigerian government should critically look into other approach to solve the power problem in Nigeria. Though, Nigerian government has tried to solve the problem by privatizing the company that supplies electricity yet since November 2013, there is no significant signal of improvement.

A regular and stable power supply could cheapen the cost of providing ICT facilities in hospitals, homes, and other public places and thus enhance the accomplishment of e-Health in Nigeria.

7.2 Government's Attitude

The Nigerian government has yet to appreciate the use of ICT in health care delivery. Government has not seen ICT as a vehicle that can drive health and provide good health care facilities for the citizenry. Presently, apart from University College Hospital Ibadan Abuja National Hospital, no teaching hospitals or medical centers have websites.

The Nigerian government should perhaps support the use of ICT in health care delivery systems by establishing an agency that will promote the deployment of ICT

in Nigerian hospitals, with a separate remit and budget from the health ministry. This should ease delivery of health care by adopting some of the benefits highlighted in this chapter.

7.3 Cost of ICT Peripherals

The price of computer hardware and software in Nigeria is very high compared with the income of an average Nigerian making them prohibitive for most people, and even government establishments, to buy. Government should encourage indigenous computer companies to produce major components in house by giving interest free loans to them. Research institutes can be used in production of the components because in developed nations like UK research institutes are the ones spearheading technological innovations and development. Nigeria government should encourage research institutes to become actively involved in software development and identification of best practice from other countries; particularly extensive, open source applications.

7.4 Telecommunication Facilities

One of the ICT development problems in Nigeria is the inadequate telecommunication facilities. Though the International Telecommunication Union (ITU) has rated Nigerian's telecommunication sector as the fastest growing technology in Africa. This applies chiefly to mobile technology, and the majority of Nigerians have no access to landlines. The Nigerian government should find means of restoring an effective and efficient telephone system that is landline, which is more economical to maintain. Also there is a need for Intranet facilities within the hospitals which will aid communication as discussed in the chapter above.

7.5 Internet Connectivity

The Internet helps in controlling costs and more importantly in transforming the flow of information in health sector. Healthcare organizations use Internet for business processes due to cost reduction which was estimated at 10:1 to 100:1 in routine transactions [28]. Communication satellite is not common in Nigeria due to the costs of equipments. As at 2003, no communication satellite was issued by NCC. Private or government-run hospitals could not host websites because of subscription and maintenance costs. On May 14, 2007, China built a communication satellite called NIGCOMSAT-1 for Nigeria, which may promote an expansion in the communication sector, though government should also consider the use of fiber optics for hospitals.

Though, the GSM operators offer Internet data plan facilities for their subscribers which can be used on laptop, iphone, ipad, and other devices. The Internet data plan is a bit expensive when compared with other developed countries of the world. Government should through NCC regulate the Internet data plan prices of these GSM operators.

7.6 Resistance to New Technology

The introduction of new technology is related to the user of the technology which may be positive and negative. Many Nigerians, like any other citizens of the world, will resist to new technological developments which will have negative impact on their job. Workers expect training on how to use new technology and a corresponding increase in their income, while the organization introducing new technology expectation may be to reduce staff strength and cost of operation. Often in Nigeria, downsizing is the issue to be raised before the introduction of new technology and this always leads to resistance by the workers, because of the fear of losing their job. In order to use ICT in Nigerian hospitals, government should train hospital staff on how to use it and not lay them off and employ those with ICT skills.

7.7 Lack of Maintenance Culture

Lack of maintenance culture is another problem. Even government agencies find it difficult to maintain ICT equipment in Nigeria. Both preventive and corrective maintenance are very important for any ICT equipment. The financial plans for purchase of any equipment should encompass the maintenance of such equipment, and allowances for depreciation in value which is not the case in many organizations in Nigeria. This could be enforced by an ICT policy banning any organization from importing, supplying, and installing any ICT equipment for another organization or itself without maintenance agreement.

7.8 ICT Literacy Rate Among Health Workers

In Nigeria, the number of computer users equals to that of mobile phones in UK and with time, the number of Internet and computer users in Nigeria may equal to that of UK. In Nigeria, mobile phone has the highest number of users, followed by computer. Generally, ICT literacy level is very low in Nigeria and the health workers in Nigeria are not left out. Most of them lack some basic ICT skills and this makes it difficult for some of them to effectively use ICT based systems in the hospitals.

Nigerian government should introduce ICT courses into the medical schools through the appropriate authority. This would be cheaper than retraining health workers on the use of ICT after graduation.

Also, government should organize seminar/workshop where health workers could be trained on the use of ICT, most especially those working in the hospitals now before the introduction of ICT skills into Nigerian medical school curriculum.

8 Conclusion

Nigeria is one of the few developing nations of the world that are backward in the use of ICT in health care delivery. If Nigeria has found itself in the application of ICT to solve most of its health needs and reduce medical tourism like other developing countries like India, Rwanda, Kenya, and Uganda which have greatly explored the potentials of ICT in transforming their health systems and economies, Nigeria would not be facing some of the challenges in health care delivery.

Nigeria needs to enthrone a strategic policy for sustained utilization of ICT in health care delivery in a way to maximize its alluring and undeniable benefits in strengthening its healthcare system, improving access of Nigerians to healthcare services in the face of limited resources and acute-on-chronic shortage of medical/health manpower, improving medical education and research, improving the quality and efficiency of healthcare services, and therefore improving the health and development indices of Nigeria. If there is the introduction of ICT in health sector in the country, there would readily be available accurate health national database across the hospitals in the country where researchers and the agency in charge of Statistics can rely on.

If Nigerian government can fully utilize opportunities offered by ICT in Nigerian hospitals, most Nigerians who travel to other countries of the world to get medical treatment would have spend such money in Nigeria which would increase Gross Domestic Product and create more employment for Nigerians.

Acknowledgement I hereby acknowledge the referees, Iyinoluwa and Iyanuoluwa, my research associates.

References

1. D. Adesina, The national health insurance scheme (2009). Available from http://www.Nigeriandoctor.Com/news.php?extend.85D. Accessed 3 Oct 2013
2. A.O. Afolabi, MINPHIS: Improving patient data in a Nigerian hospital e-Government for development eHealth case study. The eGovernment for development information exchange, University of Manchester's Institute for Development Policy and Management, 2004. Available from http://www.egov4dev.org/minphis.htm. Accessed 13 Mar 2014

3. M. Anja, Information system development in developing countries, risk management and sustainability analysis in Nigeria software companies, ed. by S. Puuronen (Department of Computer Science and Information Systems, University of Jyvaskyla, Marja-Leena Tynkkynen and pekka Olsbo, (2002), pp. 27–54
4. I.L. Arodiogbu, Introducing social health insurance to solve problems of poor health sector in Nigeria, MA thesis, University of Leeds, 2005
5. O.J. Awolola, B.L. Lawson, The creation of a national telemedicine/ehealth association in Nigeria and its impact on the Federal Ministry of Health ICT-committee (2006). Available from http://www.medetel.lu/download/2006/parallel_sessions/presentation/0407/Awolola.pdf. Accessed 17 Jan 2014
6. B. Bakare, Addressing youth unemployment in Nigeria, BusinessDay (2013). Available from http://businessdayonline.com/2013/11/addressing-youth-unemployement-in-nigeria/. Accessed 17 Feb 2014
7. I.S. Bello, F.A. Arogundade, A.A. Sanusi, I.T. Ezeoma, E.A. Abioye-Kuteyi, A. Akinsola, Knowledge and utilization of information technology among health care professionals and students in Ile-Ife, Nigeria: a case study of a university teaching hospital. J. Med. Internet Res. **6**(4), 45–49 (2004)
8. J.B. Braa, C. Nermunkh, Health information system in Mongolia: a difficult process of change, in *Information Technology in Context, Perspective from Developing Countries*, ed. by C. Avgeuru, G. Walsham (Ashgate, Burlington, 2000)
9. O.A. Daini, M. Korpela, J.O. Ojo, H.A. Soriyan, The computer in a Nigeria teaching hospital: first year experience, in *Proceedings of the MEDINFO 92*, ed. by K.C. Lun, P. Degoulet, T.E. Piemme, O. Rienhoff (Elsevier, Amsterdam, 1992)
10. J.O. Emadoye, Information communication & technologies impact on governance, social and economic development, workshop and exhibition on windows of investment opportunities in the petroleum, power & telecommunications sectors of Delta State, held at Nelrose Hotel, Asaba, 2002
11. F. Feliciani, Medical care from space: telemedicine. ESA Bull. **114**, 54–9 (2003)
12. HELINA, First international working conference on health informatics in Africa [online] (1993). Available from http://www.africa.upenn.edu/Listserv/Health_informatics_15163.html. Accessed 17 July 2007
13. D. Hyman, C. Silver, The poor state of health care quality in the US: IS malpractice liability part of the problem or part of the solution. *Research Paper Series* [online] (2004). Available from http://papers.ssrn.com/sol3/. Accessed 27 Oct 2013
14. P.A. Idowu, A.S. Ajayi, GSM based referral system for primary health care centers in Nigeria. Int. J. Soft Comput. **3**(6), 421–427 (2008)
15. B. Idowu, E. Ogunbodede, B. Idowu, Information and communication technology in Nigeria: the health sector experience. J. Inf. Technol. Impact **3**(2), 69–76 (2003)
16. P.A. Idowu, E.R. Adagunodo, E.O. Ogunbodede, Distributed resource sharing system for hospitals in Nigeria. J. Biomed. Soft Comput. Hum. Sci. **9**(2), 1–7 (2003)
17. P.A. Idowu, R.E. Adagunodo, R.A. Adedoyin, Information technology infusion model for health sector in a developing country: Nigeria as a case. Technol. Health Care **14**(2), 69–77 (2006)
18. P.A. Idowu, R.E. Adagunodo, S. Olateju, Radiological information system for primary health centres in Nigeria. J. Comput. Sci. Appl. **12**(1), 80–90 (2006)
19. P.A. Idowu, E.A. Olajubu, G.A. Aderounmu, E.R. Adagunodo, Wireless data processing model in hospital environment: a case study of Obafemi Awolowo University teaching hospital. Biomed. Fuzzy Hum. Sci. **12**(1), 23–27 (2007)
20. P. Idowu, D. Cornford, L. Bastin, Health informatics development in Nigeria. J. Health Inform. Dev. Ctries. **2**(1), 15–23 (2008)
21. P.A. Idowu, N. Okonrokwo, E.R. Adagunodo, Spatial predictive model for malaria in Nigeria. J. Health Inform. Dev. Ctries. **3**(3), 9–17 (2009)
22. P.A. Idowu, E.R. Adagunodo, O.A. Esimai, Development of a web based environmental health tracking system for Nigeria. Int. J. Inf. Technol. Comput. Sci. **7**(4), 61–71 (2012)

23. P.A. Idowu, E.R. Adagunodo, O.A. Esimai, T.C. Olapade, Development ocf a web based GIS waste disposal management system for Nigeria. Int. J. Inf. Eng. Electron. Bus. **3**(4), 40–48 (2012)
24. L. Kannegiesser, Nation health insurance scheme to boost generics market in Nigeria (2009). Available from http://www.frost.com/prod/servlet/market-insihttp.pag?Src=RSS&docid=155485216. Accessed 7 Mar 2014
25. M.R. Myers, Telemedicine: an emerging health care technology. Health Care Manag. **22**(3), 219–23 (2003)
26. NHIS, National Health Insurance Scheme Decree No. 35 of 1999 (1999). Available from http://www.nigeria-law.org/NationalHealthInsuranceSchemeDegree.htm.. Accessed 27 Dec 2013
27. NHS, *Information for Health: An Information Strategy for the Modern NHS, 1998–2005*. (NHS, Leeds, 1998)
28. P.W. O'Carroll, W.A. Yasnoff, M.E. Ward, L.H. Ripp, E.L. Martin (eds.), *Public Health Informatics and Information Systems* (Springer, New York, 2003)
29. A.O. Odusote, ICT for public health care delivery in Nigeria: challenges, opportunities and milestones. Paper presented at the E-Nigeria 2010 conference, 2010
30. A. Omotosho, O.J. Emuoyibofarhe, O. Adegboola, ICT in health care delivery system: a framework for developing nations (2012). Available online from http://www.wikehealth.org/sites/default/files/whitepaper/139/ICTforHealthPaper.pdf. Accessed 22 Apr 2014
31. D. Remlex, Information and communication technology in chronic disease care. Med. Care Res. Rev. **64**(2), 123–147 (2007)
32. F. Shiferaw, M. Zolfo, The role of ICT towards universal health coverage: the first steps of a telemedicine project in Ethiopia. Glob. Health Action **5**, 1–8 (2012)
33. H.R. Warner, D.K. Sorenson, O. Bouhaddou, The expert system model, in *Computers and Medicine: Knowledge Engineering in Health Informatics* (Springer, New York, 1997)
34. WOCON, Nigeria go report of the decade of review of the implementation of Beijing platform for action (2004). Available from http://www.wildaf-ao.org/eng/IMG/doc/Nigeria_ENG-2.doc. Accessed 27 July 2013
35. World Health Organisation, Health-related millennium development goals, world health statistics. WHO library cataloguing-in-publication data (2009). ISBN:9789241563819. Available from http://www.who.int/whosis/whostat/EN_WHS09_Full.pdf. Accessed 12 Mar 2014

Part II
On Social Networks and Economical Aspects

Community Analysis and Link Prediction in Dynamic Social Networks

Blaise Ngonmang, Emmanuel Viennet, Maurice Tchuente, and Vanessa Kamga

Abstract Community detection and link prediction are two well-studied problems in social network analysis. They are interesting because they can be used as building blocks for other more complex problems like network visualisation or social recommendation. Because real networks are subject to constant evolution, these problems have also been extended to dynamic networks. This chapter presents an overview on these two problems.

1 Introduction

Many real-world complex systems can be modelled as networks. A network is a set of entities, called nodes or vertices, connected by links also called edges. The semantics of entities and links depends on the underlying system. For example, in social networks, entities are persons and links are social relationships such as friendship, message exchanges, or collaborations. In power grids, vertices correspond to stations and substations, and edges represent physical transmission lines.

Complex networks generated from real-world systems usually share an interesting property called community structure [15]. A community is a set of nodes having more connections between them than with the rest of the network. These communities can be interpreted as modules and can help to analyse and visualise

B. Ngonmang (✉) • V. Kamga
Université de Paris 13, Sorbonne Paris Cité, L2TI, 93430 Villetaneuse, France

Université de Yaoundé 1, IRD UMI 209 UMMISCO-LIRIMA, Equipe IDASCO, BP 812 Yaoundé, Cameroon
e-mail: blaise.ngonmang@univ-paris13.fr; vansylvania.kamga@lirima.org

E. Viennet
Université de Paris 13, Sorbonne Paris Cité, L2TI, 93430 Villetaneuse, France
e-mail: emmanuel.viennet@univ-paris13.fr

M. Tchuente
Université de Yaoundé 1, IRD UMI 209 UMMISCO-LIRIMA,
Equipe IDASCO, BP 812 Yaoundé, Cameroon
e-mail: maurice.tchuente@lirima.org

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_5

the network. Global community detection methods generally assume that the entire structure of the network is known. This assumption is not realistic for very large and dynamic networks. Moreover, these methods usually produce very large communities [16] that are not very useful in practice [40]. For that reason, the concept of local community, i.e. a community obtained by exploring the network starting from a node u_0, is considered. The methods introduced for the identification of local communities do not require to access the entire network, allowing real time processing [10, 31, 39].

Work on social networks has for a long time considered only a static view: a snapshot G_t is taken at a particular time t and is analysed. However, networks are dynamic by nature. New nodes appear and some existing nodes disappear. Similarly, links representing social relations are created or ended. This dynamics can be captured by considering T snapshots $G = (G_1, G_2, \ldots, G_T)$ of the network at times $1, 2, \ldots, T$. One can then design algorithms to predict links as well as local and global communities that are likely to appear in the next snapshot, G_{T+1}.

The remainder of this chapter is organised as follows: Sect. 2 gives some general definitions and observations on complex networks. Global community detection methods are discussed in Sect. 3. Section 4 then presents some recent methods for the detection of local communities detection and the analysis of their dynamic. Section 5 presents link prediction and some methods defined for this problem. Finally Sect. 6 presents the conclusion and draws some perspectives.

2 Complex Networks Analysis

A social network can be represented by a graph $G = (V, E)$, where V is the set of vertices or nodes, and E is the set of edges or links, formed by pairs of vertices. The two nodes u and v are the end vertices of the edge $e = (u, v)$. If the order of end vertices matters in an edge, then the graph is said to be directed otherwise, it is undirected. Links of directed graphs are denoted $e = (u, v)$. The neighbourhood $\Gamma(u)$ of a node u, is the set of nodes v such that $(u, v) \in E$. The degree of a node u is the number of its neighbours or the cardinality of $\Gamma(u)$, i.e. $degree(u) = |\Gamma(u)|$. The degree of node u will also be denoted by d_u. Given this model, all graph theoretic tools can be reused in network analysis. We recall in the rest of this section the main concepts of graph theory that will be used in this paper.

Hereafter, the number of nodes of the network will be denoted by n and the number of edges will be denoted by m. The adjacency matrix of G is an $n \times n$ boolean matrix A defined by $a_{ij} = 1$ if there is a link from i to j, and $a_{ij} = 0$ otherwise. In some applications, it is useful to model the strength of the link between i and a neighbour j, and a_{ij} is a real number. Such networks are said to be weighted.

The spectrum of a graph G is the set of eigenvalues of its adjacency matrix A. The Laplacian matrix L of a graph G is defined by $L = D - A$, where D is the diagonal matrix of order n defined by $d_{ii} = degree(i)$.

A finite path of length k in G is a sequence of edges $e_1 = (u_1, u_2), e_2 = (u_2, u_3), \ldots, e_k = (u_k, u_{k+1})$, such that two consecutive edges $e_i = (u_i, u_{i+1})$

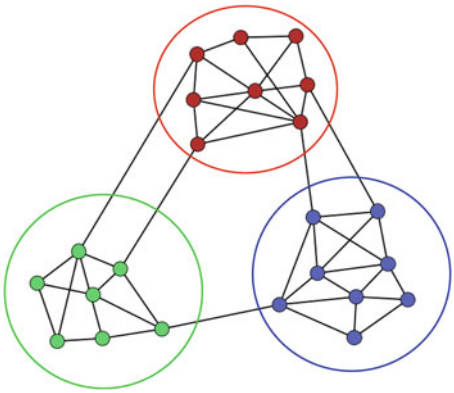

Fig. 1 Example of community structure

and $e_{i+1} = (u_{i+1}, u_{i+2})$ share a common end vertex u_{i+1}. Such a path connects u_1 to u_{k+1}. A path of length k is said to be closed if $u_1 = u_{k+1}$. A connected component of an undirected graph G is a maximal subgraph in which any two vertices are connected to each other by paths. Maximal means that such a component is not connected to any additional vertex in G. A clique is a set of nodes that forms a complete graph, i.e. with all possible links. The term $k - clique$ is used to denote a clique of k nodes.

A link $e = (i, j)$ is internal to a sub-graph $G' = (V', E')$ if i and j are in V'. A link $e = (i, j)$ is external to a sub-graph G' if either i or j is in V', but not both. The density δ of a graph corresponds to the proportion of its existing links compared to the total possible links. The internal density δ_{in} corresponds to the proportion of internal links of a sub-graph compared to the possible internal links. Similarly, the external density δ_{out} corresponds to the proportion of external links of a sub-graph compared to the possible external links.

The clustering coefficient of a network is the number of closed paths of length 3 (or triangles) divided by the number of paths of length 2. It corresponds to the probability that two nodes u and v, connected to a common neighbour w, are also connected.

A community is a set of nodes having a high internal density and a low external density. Figure 1 presents an example of community structure in a network.

It has been observed that many real-world complex networks share some characteristics [5, 35]:

- *Scale-free property:* the degree distribution follows a power law, i.e. the probability that a node has degree k is given by:

$$P(k) = k^{-\gamma} \tag{1}$$

  for a given constant γ usually between 2 and 3.
- *Small world property:* the shortest path between any given pair of nodes is usually small.
- *High clustering coefficient* compared to a random network.

- *Presence of a community structure*. Note, however, that community structure is not always present or easy to detect. This topic is the subject of active research, see, for instance, [8, 26].

The observation of a dynamic network during T time-steps is modelled by $G = (G_1, G_2, \ldots, G_T)$, where $G_t = (V_t, E_t)$ is the network observed at time t.

3 Global Communities in Social Networks

Given a network $G = (V, E)$, the global community detection problem can be defined as follows: find a partition $C = \{C_1, C_2, \ldots, C_k\}$ of nodes such that the link density is high in each C_i and low between each C_i and the rest of the network. To uncover the community structure in a network, most existing methods translate this (quite informal) definition into a computable *quality function* and then use a greedy algorithm to approximate the optimum of this function and the associated community structure. A very good survey on global community detection can be found in [15]. Note that the former definition is quite restrictive: in some cases, it is interesting to consider *overlapping* communities [21, 43].

3.1 Quality Functions

Quality functions or criteria will be used to assess how good the computed community structure is. Many quality functions have been defined in the literature. Examples are the conductance, the performance, and the modularity [15].

The *conductance* is one of the simplest functions. For a partition $S \cup \overline{S}$, it is defined as the ratio between the number of external links and $min(a(S), a(\overline{S}))$ where $a(C)$ is the total number of links having one end in C. It corresponds to the following expression:

$$\Phi(S) = \frac{|\{(u, v) : u \in S, v \in \overline{S}\}|}{min(\sum_{u \in S} d(u), \sum_{v \in \overline{S}} d(v))} \tag{2}$$

Its values range from 0 to 1. A value close to 0 corresponds to a good partition.

The *performance* [15] is the proportion of pairs of nodes correctly interpreted by the algorithm, i.e. pairs of nodes belonging to the same community and connected with links and pairs of nodes belonging to different communities and not connected. For a community structure $C = (C_1, C_2, \ldots, C_n)$, the performance score is given by:

$$P(C) = \frac{|\{(i, j) \in E, C_i = C_j\}| + |(i, j) \notin E, C_i \neq C_j|}{n(n-1)/2} \tag{3}$$

Clearly, $0 \leq P(p) \leq 1$. A value for the performance close to 1 means a good partitioning.

The *modularity* introduced by Girvan and Newman [37] is the most used quality function. The intuition behind this quality function is that a random network is not supposed to have a community structure. For each community, the internal density is compared to the expected internal density in a random network with the same number of nodes and the same degree distribution but without community structure. This corresponds to the formula:

$$Q = \frac{1}{2m} \sum_{ij} (A_{ij} - \frac{d_i d_j}{2m}) \delta(C_i, C_j) \tag{4}$$

with m the number of links of the network, A the adjacency matrix, C_i the community of node i. The function $\delta(C_i, C_j)$ is equal to 1 if $C_i = C_j$ and 0 otherwise. The value of Q ranges from -1 to 1. High values of Q are supposed to correspond to partitions with well-separated communities.

Modularity optimisation gives a good way to detect community structures in very large networks. However, this quality function has some drawbacks.

The first drawback is the *instability*. The assumption behind the modularity is that a random network is not supposed to have community structure. The actual community structure is then compared with a *null model* expressed in terms of expectation as presented in Eq. (4). This leads to many possible realisations of the null model and many (very) different community structures with a high modularity, even in random networks [27].

The second drawback is called the *resolution limit*. Indeed, it has been shown in [16] that the size of the community produced by most algorithms depends on the number of nodes in the network. It is then impossible to detect small, even well-separated communities in very large networks.

3.2 Community Detection Methods

Many global community detection methods have been proposed. The main classes of these algorithms are: random walks methods, hierarchical methods, spectral methods.

Random Walks

A random walker on $G = (V, E)$ follows a stochastic process that starts at a node $u_0 \in V$ and at each step i, selects, with probability P_j among its neighbours, the next node j to visit at time $t + 1$ [29]. Usually, this selection is done randomly

and uniformly, i.e. $P_j = \frac{a_{ij}}{d_i}$. The length of a random walk is its number of steps. The transition matrix of the random walk is $P = AD^{-1}$, and the probability of going from a vertex i to a vertex j in t steps is $(P^t)_{ij}$.

The idea behind random walk methods is that a random walker on G tends to get trapped into communities. As a consequence, two vertices i and j of the same community tend to see all the other vertices in the same way, i.e. if the length t of the random walk is long enough, the i^{th} row $(P^t)_{i.}$ and the j^{th} row $(P^t)_{j.}$ will be similar. This leads to a definition of distance between nodes and community detection becomes a clustering problem that can be solved using, for example, hierarchical methods [45].

Hierarchical Methods

Hierarchical methods are either top down or bottom up. In top down methods, one starts with all nodes in one unique community. At each step, one tries to separate the existing communities into sub-communities. An example of method ranging in this category is described in [37]. In bottom up methods, one starts with each node belonging to a separate community and at each step, one tries to merge the most similar communities.

Louvain [7] is one of the fastest methods for community detection in complex networks. It can be used in the general case of weighted networks (links have weights expressing the strength of relationships). The algorithm is a bottom up hierarchical method and has two main steps. In the first step, every node of the network is evaluated by computing the weighted modularity gain if it is then added to the community of a neighbour j. Node i is added to the community which produces the maximal positive gain. This process is repeated until there is no more positive gain. At the end of this step one has a partition of the network.

In the second step, each of the previous discovered communities becomes a *super-node*. The weight of the link between two *super-nodes* is the sum of the weights of links between nodes of the corresponding communities. Links between nodes of the same community lead to a self-loop of the corresponding *super-node*.

These two steps are repeated until no more gain in modularity is observed. The speed of this method comes from the observation that the modularity can be optimised locally: only the neighbours are considered during the evaluation. This enables to update the modularity gain in linear time ($O(m)$). This method is only limited by the storage (main memory) capacity and allows to analyse networks with millions of nodes in few minutes.

Relation with Spectral Methods

Random walks on graphs are strongly related to methods that study community structure using spectral properties of graphs. For instance, in [45] it is shown that

the distance r_{ij} defined between two nodes i and j, using random walks, is related to the spectral property of the matrix P by the formula:

$$r_{ij}^2 = \sum_{\alpha=2}^{n} \lambda_\alpha^{2t} (v_\alpha(i) - v_\alpha(j))^2 \tag{5}$$

where $(\lambda_\alpha)_{1\leq\alpha\leq n}$ and $(v_\alpha)_{1\leq\alpha\leq n}$, are respectively, the eigenvalues and right eigenvectors.

On the other hand, it is shown in [49] that if two vertices i and j belong to the same community, then the coordinates v_i and v_j are similar in eigenvectors of P that correspond to the largest positive eigenvalues. More recently, Newman has shown that community detection by *modularity maximisation*, community detection by *statistical inference*, and *normalised-cut* graph partitioning can be solved using spectral approaches that use matrix representations of the network [36].

3.3 Dynamics of Global Communities

For community detection in dynamic networks, various methods have been proposed in the literature. Some of these methods try to track the evolution of communities between time-steps, others try to update the existing community structure and finally, the last class of methods try to detect communities that are consistent in all the time steps.

Community Tracking

Algorithms for global community detection in static networks have been used for tracking community structures in dynamic networks. The idea is to detect the communities at each time-step and match them between consecutive time-steps. Examples of such methods were proposed by Palla et al. [44], Greene et al. [20], and Tantipathananandh et al. [50].

The algorithm by Palla et al. is based on the Clique Percolation Method [43]. This approach is not usable in very large graphs. The two other methods can be used with any algorithm and particularly with modularity-based algorithms.

The tracking of communities is very difficult because of the instability of global methods. This is particularly true for modularity-based methods. This drawback can be reduced by using community cores analysis. In [46] and [14] a community core is defined as a set of node that are frequently in the same community during many executions of an unstable algorithm.

Community Updating

The main idea of community updating is to detect the community structure at a reference time-step t_0 and then, for all the future time-steps, the community structure is updated according to the elementary events that can appear in the network. These elementary events are: addition of a node, deletion of a node, creation of a link, and removal of a link. Examples of methods that have been proposed to handle these events can be found in [9, 42].

The main problem with community updating is that it is sensible to the initial partition. Note that community core analysis can help to stabilise the initial partition.

Long-Term Community Detection

A long-term community can be defined as a set of nodes that interact more with each other than with the rest of the network at all time-steps [4]. This is an extension of the classical community definition to the dynamic context. To detect long-term communities, Aynaud and Guillaume [4] have proposed to define an average modularity and to optimise it using a modified version of Louvain's algorithm [7]. In another work, Mitra et al. [33] have proposed a method designed for citation networks. They propose to build a summary network as follows:

- a node a_t is created if author a has published a paper at time t.
- a link is created between nodes a_t and $b_{t'}$ if and only if the paper published by author a at time t cites the paper published by author b at time t'.

A static community detection method can then be used to mine the community structure in this summary network.

4 Dynamics of Local Communities

Global communities give a way to analyse the dynamics at a macroscopic level. Because global communities are either too large or cannot be computed due to the size of the network, we show in this section how to analyse the dynamics using local communities that can be computed without the entire knowledge of the network.

4.1 Local Community Identification

Given a node u_0 of a partially known network $G = (V, E)$, initially limited to u_0 and its neighbours, and with the restriction that new information can only be obtained by getting adjacent nodes one by one, the problem of local community identification is to find the community the node u_0 belongs to.

Table 1 Some quality functions for local community identification

Quality functions	Authors
$R = \frac{B_{in}}{B_{in}+B_{out}}$	Clauset [11]
$M = \frac{D_{in}}{D_{out}}$	Luo et al. [31]
$L = \frac{\sum_{i\in D}\frac{\lvert\Gamma(i)\cap D\rvert}{\lvert D\rvert}}{\sum_{i\in B}\frac{\lvert\Gamma(i)\cap S\rvert}{\lvert B\rvert}}$	Chen et al. [10]
$T = \frac{\sum_{i\in D}\frac{\lvert\Gamma(i)\cap D\rvert}{(1+d_i)}}{\sum_{i\in D}\lvert\Gamma(i)\cap S\rvert(1+d_i)}$	Ngonmang et al. [39]

Most existing algorithms for local community identification use a greedy scheme: initially, the local community D contains only the starting node u_0 and the quality of this initial community is 0. At each step, the external node that maximises the quality function F used by the algorithm is considered. If its inclusion into D increases the quality criterion F, then it is added to D, and the quality F of the community is updated. This procedure is repeated until there is no more external vertex whose inclusion into D increases the quality F. At the end of the algorithm, D contains the local community of u_0.

Let D denote a local community. B is the set of nodes of D that have at least one neighbour out of D and S is the set of nodes out of D that have at least one neighbour in D. Table 1 presents some existing quality functions used in local community identification. In this table, D_{in} corresponds to the set of links having both ends in D. D_{out} corresponds to the set of links having only one end into D. B_{in} and B_{out} have a similar meaning. d_i is the distance from i to u_0, the starting node, and $\Gamma(i)$ is the set of neighbours of node i.

Some problems are similar to local community identification because they only require local information. One of these problems is ego-community detection. This problem consists in detecting the communities between the direct neighbours of a node. One successful approach to solve this problem is the method proposed in [19]. Unfortunately, this method is not suitable for local community detection because it discards the rest of the network. Another similar problem is to consider multiple starting nodes and detect the communities that contain them, as in the work of Danisch et al. [12].

4.2 Application to a Dynamic Behaviour: Churn Prediction

The objective of churn prediction is to estimate the likelihood that a given user will stop using a social network platform in the near future. A churner will thus be defined as a user who has become inactive for a certain period of time. This knowledge can be exploited by the platform operator to take preventive actions: if the user is likely to stop using the platform, it could be interesting to send him some incentives (personalised recommendations, free applications, etc.).

Most of the methods for churn prediction belong to three main categories: feature-based methods, network-based methods, and hybrid methods. Feature-based methods extract attributes from the user profile (age, gender, etc.) and usage (time spent, connexion history, etc.) of the platform and then build a predictive model [24].

Network-based methods use the social links to detect the churners. The methods of this category usually model the churn prediction as a diffusion or contagion process [13]: starting with the known churners as seeds, each seed tries to activate its neighbours at each iteration. This process is repeated until convergence or up to a maximum number of iterations.

Finally hybrid methods combine the two previous ones. One application of local community analysis can be found in this category. Indeed, the hybrid method proposed in [40], for example, proposes to extract some attributes from the local community of the node and to add them to the features of the node in order to build the churn prediction model.

4.3 Prediction of Local Communities

The local community prediction problem in complex networks can be stated as follows: given a dynamic network $G = (G_1, \ldots, G_T)$ and the dynamic local community $D = (D_1, \ldots, D_T)$ of a node u_0, what will be the local community D_{T+1} at time $T + 1$ in (G_{T+1})?

To solve this problem, two main classes of approaches can be used. The first class consists in predicting for each node v, whether or not it will belong to the local community D_{T+1} of node u_0 at time $T + 1$. The second class first predicts the structure of the network at time $T + 1$ and then computes the local community of u_0 in the predicted network.

To predict the membership of each node u to D_{T+1}, some simple attributes can be computed at each time-step: the position of u with respect to the subsets D, B, and S defined in Sect. 4.1, the number of links with the community, the number of links with nodes outside the community, etc. A supervised learning model can then be used for the prediction. The real challenge here is that one is restricted to nodes having already belonged to the local community of u_0 or to its neighbourhood.

The work in [41], for example, presents a method to predict local communities according to the second approach. The network of the target time-step is predicted and the local communities are computed on that predicted network. To keep the locality on this process, only the local portion of the network containing the starting node is predicted.

This gives a way to locally analyse a dynamic network and make some predictions on its future structure. A more microscopic view of the dynamic is provided by link prediction that will be presented in the next section.

5 Link Prediction

Given a snapshot of a social network at time t, the link prediction problem is to accurately predict the edges that will be added to the network from time t to a given future time t'. The link prediction problem therefore tackles the following question: to what extent can the evolution of a social network be modelled using features intrinsic to the network itself [28]? Formally, consider a network $G = (V, E)$ where V is the set of vertices and E is the set of links. The set of edges $(u, v) \subseteq V$ with $u \neq v$ that are absent in E is denoted $\overline{E}$. In a practical application, $\overline{E}$ can be divided into two parts: the set E' of links that will appear in the future, also called missing links, and the set E'' of edges that will never appear. Clearly, $E' \cup E'' = \overline{E}$ and $E' \cap E'' = \emptyset$. The challenge of link prediction is to produce quickly, accurate approximations for E', even for huge social networks.

As noted by Zhu and Kinzel [53], it is possible, for any discrete prediction algorithm A for sequences, to generate using an algorithm B no more complicated than A, an instance for which A's prediction is always wrong. Moreover, for any prediction algorithm A and an instance x, there exists a sequence y no more complicated than x, such that if A performs better than random on x, then it will perform worse than random on y by the same margin. This shows that, to design a predictor with good performance, it is necessary to have prior knowledge on the problem.

Link prediction is a very active research area because of its wide range of applications. For instance, if G is a social network representing recorded interactions between terrorists, the link prediction can be used to detect underground relationships between them. On the other hand, a link prediction algorithm can be applied to a clients/products network produced by an e-commerce platform, to suggest products that a client is likely to purchase in the near future. Other algorithms and applications related to link prediction in complex networks can be found in [30].

We now present some basic link prediction algorithms according to the following nomenclature: probabilistic methods, transitivity-based methods, and attributes-based methods. After that we introduce an extension of the link prediction problem to dynamic networks.

5.1 Probabilistic Methods

The most naive probabilistic model of link prediction is the Random predictor which randomly chooses a subset of links that are not present in the network and predicts them. Since the subset selection is done randomly, the accuracy of the algorithm is based on luck. The probability of failure of the Random predictor is $\frac{1}{2}$. This method can't be taken seriously when dealing with an application. It only serves as reference point: any serious algorithm must have a better accuracy.

The probabilistic approaches can nevertheless be useful when there is a prior knowledge on the problem. For instance, many complex natural and social systems assemble and evolve through the addition and removal of nodes and links. This dynamics often appears to be a self-organising mechanism governed by evolutionary laws that lead to some common topological features. One of such features is the power-law degree distribution, i.e. the probability that a node has degree k is $P(k) = k^{-\gamma}$, usually with $2 < \gamma \leq 3$. Such networks are said to be scale-free. For such networks, the "preferential attachment principle" states as follows: when a new node is added to the network with m edges that link this new node to m nodes already present, the probability that this new node will connect to a node i with degree d_i is proportional to d_i, i.e. $\pi(d_i) = \frac{d_i}{(\sum_i d_i)}$. It can be shown that a network evolving according to this principle tends to a scale-invariant state with $\gamma = 3$. Clearly, such a model of network growth constitutes an a priori information that can help to design efficient link prediction algorithms. The preferential attachment principle is also known in economy as cumulative advantage: the rich get richer [6, 48].

The preferential attachment is a good illustration of Zhu and Kinzel's observation. Indeed, it gives the worst performance when applied to physical Internet networks where high degree nodes are routers that have a very low probability of being connected by new physical lines.

Recently, Freno et al. [17] have proposed a new approach that is not based on parametric assumptions concerning the modelled distributions. More precisely, they have introduced the Fielder random field model, called Fielder delta statistic that, for each binary edge variable $X_{u,v}$, defines a potential that encapsulates the measure of its role in determining the connectivity of its neighbourhood. The trick is that these potentials can be estimated from data by minimising a suitable objective function. Experiments on some real-world networks have resulted in link prediction algorithms that outperform the solutions proposed by Watts-Strogatz [51] and Barabasi-Albert [6]. Other probabilistic methods for link prediction are reported in [30].

5.2 Transitivity-Based Methods

In mathematics, a binary relation $\Re$ defined on a domain D is said to be transitive if whenever u is in relation with $v(u \Re v)$ and v is in relation with $w(v \Re w)$, then u is in relation with $w(u \Re w)$.

In topological transitivity applied to a complex network $G = (V, E)$, the domain D consists of the set V of nodes of the network, and the relation $\Re$ is represented by the set E of edges. The application of topological transitivity to link prediction is based on the assumption that, as a complex network evolves, it tends to become transitive, i.e.: if at time t, u is related to v and v is related to w, then there is a high probability that at a future time t', u will be related to w. This assumption follows

from a common observation made, for instance, on friendship networks: a friend of your friend is likely to be or become your friend. This corresponds to triangles in G, i.e. triples of edges (u, v), (v, w), and (u, w). In graph-theoretic terms, the degree of transitivity of a network G can be measured by the so-called clustering coefficient [51]:

$$C = \frac{\sum_{u \in V} C_u}{|V|} \tag{6}$$

where

$$C_u = \frac{number\ of\ triangles\ connected\ to\ vertex\ u}{number\ of\ triples\ centred\ on\ vertex\ u} \tag{7}$$

As reported in [35], this coefficient has remarkable values for many current networks: greater than 0.75 for film actors and power grids; between 0.6 and 0.74 for biology co-authorship, train routes and metabolic networks; between 0.30 and 0.59 for math co-authorship, Internet and word co-occurrences in web pages, and less than 0.20 for email messages and freshwater food web.

The basic link prediction methods based on topological transitivity use some local or global properties of the network G, to assign a connection weight $Score(u, v)$, to pairs of nodes (u, v) of V. All non-observed links are then ranked in decreasing order of $Score(u, v)$. In this approach, links with the highest scores are supposed to be of higher existence likelihoods and are produced by the algorithm. Such a measure must reflect the proximity or similarity between nodes u and v. The problem is to design good similarity measures.

Let us denote $\Gamma(u)$ the set of neighbours of node u, and let $|A|$ be the cardinality of a set A. $CN(u, v) = |\Gamma(u) \cap \Gamma(v)|$ [34] corresponds to the number of common neighbours of u and v. The idea is that if u and v have many neighbours in common, then there is a high probability that they will become neighbours in the future. The efficiency of this measure has been experienced with collaborative networks [35]. However, this measure suffers from serious drawbacks. For instance, in a friendship network, the fact that two nodes u and w have a common very popular neighbour v, i.e. with a very high degree d_v, does not necessarily mean that u and w will become friends in the future. They may even be from different continents and never meet. In the same way, in an allocation network, if node u sends a unit of resource to a very popular neighbour v that serves as intermediary, and if node v subdivides the resource and sends equal parts to his neighbours, then the portion received by any neighbour $w \in \Gamma(v)$ will be $\frac{1}{d_v}$. This means that the contribution of an intermediate node v for the "future connection" between u and w is divided by the degree of v. This has motived some authors to introduce $RA(u, v) = \sum_{w \in \Gamma(u) \cap \Gamma(v)} \frac{1}{d_w}$ [52] and the log form $AA(u, v) = \sum_{w \in \Gamma(u) \cap \Gamma(v)} \frac{1}{\log(d_w)}$ [1]. Many other variants of CN have been proposed, but extensive experiments on real-world networks have shown that RA is the best whereas CN is the second best.

A nice link prediction method based on topological transitivity has been introduced by Latapy et al. [2]. Consider a bipartite clients/products network

$G = (\perp, \top, E)$ where $\perp$ is the set of clients, $\top$ the set of products, and E the set of purchases. The $\perp$-projection of G is the graph $G_\perp = (\perp, E_\perp)$ in which $(u, v) \in E_\perp$ if u and v have at least s neighbours in common in G, where s is a given threshold, i.e. $|\Gamma(u) \cap \Gamma(v)| \geq s$. The underlying intuition of the internal link prediction method is that, in a clients/products network, if two clients have bought in the past many common products, then they will probably acquire new common products in the future. This method falls within the transitivity framework as follows: if client A is related to client B in $G_\perp$ and if client B is related to product p in G, then there is high probability for A to be related to p in the future.

Another topological transitivity measure for link prediction is based on random walks already introduced in section "Random Walks" for community detection. In the simplest version of this method, it is assumed that, when a random walker is at node u, it can go in one step to any node $v \in \Gamma(u)$ with probability $\frac{1}{d_u}$. Let $m(u, v)$ denote the average number of steps necessary for a random walker to go from u to v. The commute time is the symmetrical measure $CT(u, v) = m(u, v) + m(v, u)$. This transitivity measure is then used to predict missing links: the smaller $CT(u, v)$ is, the greater is the probability for u and v to establish a connection in the future.

Association rules originally defined for large databases of sales transactions can be adapted for link prediction on a network $G = (V, E)$. Consider $D = \{\Gamma(u) | u \in V\}$. Define frequent groups of nodes as subsets that are included in at least s elements of D, where s is a given threshold. An association rule is an implication of the form $A \rightarrow B$, where $A \subseteq V$, $B \subseteq V$, $A \cap B = \emptyset$, and $A \cup B$ is frequent. A rule $A \rightarrow B$ holds with confidence c if $c\%$ of neighbourhoods in D that contain A also contain B. Hereafter, we denote $A \rightarrow B : c$. The transitivity principle states as follows: if $A \rightarrow B : c$ and $B \rightarrow C : c'$, then $A \rightarrow C : c \times c'$. In the context of an application to co-authorship [25]: A, B and C are sets of co-authors. $A \rightarrow B : c$ means that $c\%$ of articles co-authored by A are also co-authored by B, and $B \rightarrow C : c'$ means that $c'\%$ of articles co-authored by B are also co-authored by C. As a consequence, if $A \rightarrow B : c$ and $B \rightarrow C : c'$ are observed, then $A \rightarrow C$ is predicted with probability $c \times c'$ (i.e. a new article with $A \cup C$ as co-authors is predicted) [25].

5.3 Attributes-Based Methods

The great specificity for graphs that model social networks is that nodes and links usually have attributes. Consider a phone network in which a node represents a person and each link represents a call. Phone numbers can be used as node attributes and the average number of calls between nodes can be used as link attributes.

The link prediction problem can be expressed as a classification problem for pairs (u, v). The following attributes may be considered when dealing with co-authorship networks [23]: the number of common neighbours $(\Gamma(u) \cap \Gamma(v))$, the number of common keywords $(Kw(u) \cap Kw(v))$, or the total number of articles published by u and v. The class attribute is a binary variable with value 1 if the link will

appear and 0 otherwise. All attributes values are normalised to have zero mean and one standard deviation. A classification model such as Decisions Tree (DT), Support Vector Machine (SVM) or Artificial Neural Network (ANN) can then be used. Hasan et al. [23] have shown on two networks (DBPL and BIOBASE) that SVM beats all the most used classification methods.

The similarity between two nodes can use attributes of nodes and links. This is the case for *Abstract* proposed in [25], which takes into account summaries of articles in the bipartite graph Authors/Articles. The idea is that articles already published contain information on topics that interest the co-authors. It is then natural to suppose that authors working in the same domain are more likely to collaborate and co-publish an article in the future. The attributes-based similarity between two u and v authors is then defined as:

$$score(u, v) = cos(V(u), V(v)) \tag{8}$$

where $V(u)$ is a descriptor that encapsulates the attributes for vertex u. It has been shown in [25] that this approach produces very good predictions for some well-known co-authorship networks.

5.4 Link Prediction in Dynamic Complex Networks

Classical link prediction is not sufficient to fully analyse the dynamics of a complex network. Indeed, it supposes that all the created links will last forever. However, in a real interaction complex network, the links between nodes are created and ended. For example, in a collaboration network, a publication between two scientists in a particular year does not guarantee that they will still work together in the future. To fully analyse the dynamics of connections between nodes, one needs a more general model which determines whether or not a particular link will exist at a future time t'.

This can be stated as follows: given a dynamic network $G = (G_1, \ldots, G_T)$, what will be the structure of the following snapshot (G_{T+1})? This problem is a generalisation of the link prediction problem: here not only the non-previously seen links are predicted but also the existing ones to check whether or not they will still exist in the following snapshot.

As a generalisation of the link prediction problem, the class of methods designed for the classical link prediction problem can also be used to solve it. A similarity-based and a supervised learning method to solve this problem can be found in [38]. The supervised method can be described as follows: for each snapshot t of the training period, the following features are computed for each pair of nodes:

- the number of common neighbours
- the number of common community members
- a boolean attribute indicating whether an interaction is present or not between the two nodes
- other similarity between the two nodes (if available)

The real classes are obtained on the test period. It is worth nothing that to reduce the complexity (the number of possible interactions is in the order of $O(n^2)$), only the interactions that are likely to appear, based on the computed similarity scores (topological or attribute based), are considered. Finally, a supervised learning method (SVMs for example) can be used to build the model.

6 Conclusions and Perspectives

In this chapter, we have presented some tools for the analysis of the community structure and for link prediction in social networks in both static and dynamic contexts. This corresponds to the observation of the evolution of such a network at the macroscopic level (global communities), intermediate level (local communities), and microscopic (link prediction) level. Such analysis, and more generally social mining techniques, can lead to many applications in Africa and some interesting work has already been done.

In epidemiology, for example, the work in [18] presents an agent-based model of epidemic spread using social networks. On the other hand, because of the dense interaction patterns, infectious diseases tend to spread more rapidly in communities. As a consequence, an interesting question may be to design models of infection spread, that take into account the community structure of a contact network. An example of attempt in this direction can be found in [3] where the Ross–Macdonald model which describes the dynamics of malaria has been extended to multipatch systems. In this approach, a patch models a community. On the other hand, a nice and simple analytic formula of the basic reproduction number has been proposed in [47] for cellular SIR networks.

In telecommunication networks, thanks to the Data for Development challenge http://www.fr.d4d.orange.com/, many researches have been conducted on the Ivory Coast telecommunication network [6]. For example, the work in [20] presents a method based on the phone calls network and airtime credit, for the evaluation of the socioeconomic state of a country. More recently in a Ph.D. thesis, Guigourès [22] has studied the co-clustering technique that consists of simultaneously partitioning the rows and the columns of a data matrix. Applications to detailed call records from a telecom operator in Ivory Coast have permitted to detect individuals that are the most representative of their profiles. This information was then used to improve the knowledge of users, develop new products and improve urban infrastructure related to mobility.

For co-authorship networks, the work in [32] analyses the evolution of co-publications in the community of researchers involved since 1992 in the African conference for research in applied mathematics and computer science (CARI).

Note that, for Africa to take full benefit of tools for network analysis, an increased effort must be put on data collection.

References

1. L.A. Adamic, E. Adar, Friends and neighbors on the web. Soc. Netw. **25**, 211–230 (2003)
2. O. Allali, C. Magnien, M. Latapy, Link prediction in bipartite graphs using internal links and weighted projection, in *Proceedings of the Third International Workshop on Network Science for Communication Networks (NetSci- Com)* (2011)
3. P. Auger, E. Kouokam, G. Sallet, M. Tchuente, B. Tsanou, The Ross–Macdonald model in a patchy environment. Math. Biosci. **216**(2), 123–131 (2008)
4. T. Aynaud, J.L. Guillaume, Static community detection algorithms for evolving networks, in *WiOpt'10: Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Avignon (2010), pp. 508–514. http://hal.inria.fr/inria-00492058
5. A.L. Barabasi, *Linked: How Everything is Connected to Everything Else and What It Means*, reissue edn. Plume, (2003)
6. A.L. Barabási, R. Albert, Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
7. V.D. Blondel, J. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. **10** 10008 (2008)
8. R. Campigotto, J.L. Guillaume, M. Seifi, The power of consensus: random graphs have no communities, in *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (ACM, New York, 2013), pp. 272–276
9. R. Cazabet, F. Amblard, Simulate to detect: a multi-agent system for community detection. IAT, 402–408. IEEE Computer Society (2011)
10. J. Chen, O.R. Zaiane, R. Goebel, Local communities identification in social networks, in *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'09)* (2009), pp. 237–242
11. A. Clauset, Finding local community structure in networks. Phys. Rev. **72**, 026132 (2005)
12. M. Danisch, J.L. Guillaume, B.L. Grand, Towards multi-ego-centred communities: a node similarity approach. J. Web Based Communities **9**(3), 299–322 (2013)
13. K. Dasgupta, R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjea, A. Nanavati, A. Joshi, Social ties and their relevance to churn in mobile telecom networks, in *EDBT '08: Proceedings of the 11th International Conference on Extending Database Technology* (2008), pp. 668–677
14. V. De Leo, G. Santoboni, F. Cerina, M. Mureddu, L. Secchi, A. Chessa, Community core detection in transportation networks. Phys. Rev. E **88**(4), 042810 (2013)
15. S. Fortunato, Community detection in graphs. Phys. Rep. **486**, 75–174 (2010)
16. S. Fortunato, M. Barthélemy, Resolution limit in community detection. Proc. Natl. Acad. Sci. **104**(1), 36–41 (2007)
17. A. Freno, C. Garriga Gemma, M. Keller, Learning to recommend links using graph structure and node content, in *Neural Information Processing Systems Workshop on Choice Models and Preference Learning* (2011)
18. E. Frías-Martínez, G. Williamson, V. Frías-Martínez, An agent-based model of epidemic spread using human mobility and social network information, in *SocialCom/PASSAT* (IEEE), pp. 57–64
19. A. Friggeri, G. Chelius, E. Fleury, Egomunities, exploring socially cohesive person-based communities. CoRR. abs/1102.2623 (2011)
20. D. Greene, D. Doyle, P. Cunningham, Tracking the evolution of communities in dynamic social networks, in *Proceedings of the 2010 International Conference on Advances in Social Networks Analysis and Mining, ASONAM '10* (IEEE Computer Society, Washington, 2010), pp. 176–183
21. S. Gregory, Finding overlapping communities using disjoint community detection algorithms, in *Complex Networks* (2009), pp. 47–61
22. R. Guigourès, Utilisation des modèles de co-clustering pour l'analyse exploratoire des données. Ph.D. thesis in Applied mathematics, University of Paris 1 Panthéon Sorbonne
23. M.A. Hasan, V. Chaoji, S. Salem, M. Zaki, Link prediction using supervised learning, in *Proceedings of SDM 06 Workshop on Link Analysis, Counterterrorism and Security* (2006)

24. H. Hwang, T. Jung, E. Suh, An ltv model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. Expert Syst. Appl. **26**(2), 181–188 (2004)
25. V. Kamga, M. Tchuente, E. Viennet, Prévision de liens dans les graphes bipartites avec attributs. Revue des Nouvelles Technologies de l'Information (RNTI-A6) (2013)
26. I. Keller, E. Viennet, A characterization of the modular structure of complex networks based on consensual communities, in *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS)* (IEEE, Naples, Italy 2012), pp. 717–724
27. I. Keller, E. Viennet, A characterization of the modular structure of complex networks based on consensual communities, in *2013 International Conference on Signal-Image Technology Internet-Based Systems (SITIS)* (2012), pp. 717–724
28. D. Liben-Nowell, J. Kleinberg, The link prediction problem for social networks, in *Proceedings of the Twelfth International Conference on Information and Knowledge Management, CIKM '03* (ACM, New York, 2003), pp. 556–559
29. L. Lovász, Random walks on graphs: a survey. In *Combinatorics, Paul Erdös is Eighty*, eds. by D. Miklás, V.T. Sás, T. Szönyi (János Bolyai Mathematical Society, Budapest, 1996), pp. 353–398
30. L. Lu, T. Zhou, Link prediction in complex networks: a survey. Phys. A Stat. Mech. Appl. **390**(6), 1150–1170 (2011)
31. F. Luo, J.Z. Wang, E. Promislow, Exploring local community structure in large networks, in *WI'06* (2006), pp. 233–239
32. G.R. Meleu, P. Melatagia, Analyse et modélisation du cari: croissance de la communauté de chercheurs du cari, in *Conférence de Recheche en Informatique(CRI'2013), Yaoundé* (2013), pp. 83–87
33. B. Mitra, L. Tabourier, C. Roth, Intrinsically dynamic network communities. CoRR. abs/1111.2018 (2011)
34. M.E.J. Newman, Clustering and preferential attachment in growing networks. Phys. Rev. E **64**, 025102 (2001)
35. M.E.J. Newman, The structure and function of complex networks. SIAM Rev. **45**, 167–256 (2003)
36. M.E.J. Newman, Spectral methods for network community detection and graph partitioning. Phys. Rev. **884**, 042822 (2013)
37. M. Newman, M. Girvan, Community structure in social and biological networks. Proc. Natl. Acad. Sci **99**, 7821–7826 (2002)
38. B. Ngonmang, E. Viennet, Toward community dynamic through interactions prediction in complex networks, in *2013 International Conference on Signal-Image Technology Internet-Based Systems (SITIS)* (2013), pp. 462–469
39. B. Ngonmang, M. Tchuente, E. Viennet, Local communities identification in social networks. Parallel Process. Lett. **22**(1) (2012)
40. B. Ngonmang, E. Viennet, M. Tchuente, Churn prediction in a real online social network using local community analysis, in *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'12)* (2012), pp. 282–290
41. B. Ngonmang, S. Sean, R. Kirche, Monetization and services on a real online social network using social network analysis, in *2013 IEEE 13th International Conference on Data Mining Workshops* (2013), pp. 185–193
42. N. Nguyen, T. Dinh, Y. Xuan, M. Thai, Adaptive algorithms for detecting community structure in dynamic social networks, in *2011 Proceedings IEEE INFOCOM* (2011), pp. 2282–2290
43. G. Palla, I. Derényi, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society. Nature **435**, 814–818 (2005)
44. G. Palla, A.-L. Barabási, T. Vicsek, Quantifying social group evolution. Nature **446**(7136), 664–667 (2007)
45. P. Pons, M. Latapy, Computing communities in large networks using random walks. J. Graph Algorithms Appl. **10**(2), 191–218 (2006)

46. M. Seifi, I. Junier, J.B. Rouquier, S. Iskrov, J.L. Guillaume, Stable community cores in complex networks, in *Complex Networks*. Studies in Computational Intelligence, vol. 424 (Springer, Berlin/Heidelberg, 2013), pp. 87–98
47. A. Sidiki, M. Tchuente, An analytical formula for the basic reproduction number on cellular sir networks, in *Actes du Colloque Africain de Recherche en Informatique*, (2012)
48. H.A. Simon, On a class of skew distribution functions. Biometrika **42**, 198–216 (1955)
49. I. Simonsen, K.A. Eriksen, S. Maslov, K. Sneppen, Diffusion on complex networks: a way to probe their large-scale topological structures. Phys. A Stat. Theor. Phys. **336**(1–2), 163–173 (2004)
50. C. Tantipathananandh, T. Berger-Wolf, D. Kempe, A framework for community identification in dynamic social networks, in *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '07* (ACM, New York, 2007), pp. 717–726
51. D. Watts, S. Strogatz, Collective dynamics of 'small-world' networks. Nature **393**, 440–442 (1998)
52. T. Zhou, L. Lü, Y.C. Zhang, Predicting missing links via local information. Eur. Phys. J. B **71**, 623 (2009)
53. H. Zhu, W. Kinzel, Antipredictable sequences: harder to predict than random sequences. Neural Comput. **10**, 2219–2230 (1998)

Web 2.0 Computing and Social Media as Solution Enablers for Economic Development in Africa

Patricia E. Nalwoga Lutu

Abstract Web 2.0 technologies have been used to create many social media applications that have revolutionalised the way people communicate and socialise. Today, it can be claimed that citizens, especially the young generation, of every country of the world have contributed to the widespread usage of social media. Examples of social media are Facebook, Twitter, LinkedIn and YouTube. Businesses, decision makers in the public sector, news networks and other types of news media have all been quick to realise the potential benefits of social media. Social media analytics are used by many organisations to establish the sentiment of customers towards the products and services that they sell or provide. The purpose of this chapter is to highlight (1) the available and widely used social media and their adoption on the African continent, (2) organisations on the African continent that have adopted these social media, (3) how organisations that use social media can benefit from such usage, and (4) the challenges posed by the adoption and usage of social media. It is argued that, for African countries, the major benefits of using social media are the low costs of engaging with the citizens and business customers as well as the ability to analyse the effectiveness of such engagements.

1 Introduction

Social media can be defined as Internet-based media that are created and shared by communities. Web 2.0 technologies can be defined as Internet technologies which connect people and enable the sharing of media. In the last 10 years, Web 2.0 technologies and social media have revolutionalised the way people communicate and socialise. Social networks are one category of social media that facilitates the formation of communities and sharing of content such as microblogging (Twitter), meeting people and sharing content (Facebook, Twitter, LinkedIn, YouTube). Twitter, Facebook, LinkedIn and YouTube have become very popular in many countries including African countries, especially with the young generation.

P.E.N. Lutu (✉)
Department of Computer Science, University of Pretoria, Pretoria, South Africa
e-mail: Patricia.Lutu@up.ac.za

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_6

Organisations in the public and private sectors have been quick to realise the value of social media. At the present time, business, government and non-governmental organisations typically participate in social media. The reasons for participation include assessing the opinions of the public about products and services that the organisations provide, soliciting opinions from the public, and communication and collaboration between stakeholders. Big and small business organisations are routinely using social media for marketing and branding purposes [4, 20, 29]. Many government departments and other public sector organisations use social media for purposes of engaging with citizens (e.g. [26]). Development organisations such as non-government organisations (NGOs) routinely use social media to network with stakeholders and to reach and engage with developing country communities via mobile phones.

Many online tools are available for the analysis of social media data for purposes of assessing the effectiveness of social media usage. Some tools are specific to a particular service (e.g. TwitterSentiment) while others are general purpose tools that can be used to analyse data from different types of social media services [33]. Some analytics tools may be used for free while others require payment for the services. Organisations can use these online tools to build their online presence and grow their organisation's reach online. Organisations are not limited to the use of online tools for social media analytics. In fact, data can be downloaded from a social network website, stored in special databases and then analysed offline using statistical, data mining, and machine learning methods [2]. NoSQL databases are a good candidate for the storage of large quantities of unstructured data. Some of the methods of knowledge discovery and data mining that can be applied to social network data are sentiment analysis, time series analysis, and graph mining [2, 24, 34].

Given the foregoing discussion, it is argued in this chapter that Web 2.0 technologies and social media may be viewed as a key solution enablers for government and NGOs, and private sector organisations in African countries. These technologies have the potential to connect the citizens, public sector organisations, private sector organisations and development organisations on the African continent at a very low cost and in very short time frames. The objectives of this chapter are to highlight (1) the available and widely used social media and their adoption on the African continent, (2) organisations on the African continent that have adopted these social media, (3) how organisations that use social media can benefit from such usage, and (4) the challenges posed by the adoption and usage of social media. The rest of this chapter is organised as follows: Section 2 provides a discussion of Web 2.0 and social media. Section 3 provides a discussion of organisations that use social media. Section 4 discusses social media adoption in African countries. Section 5 discusses online social media analytics tools. Section 6 discusses offline analysis of social media data. Section 7 provides a discussion of the benefits, challenges and some recommendations for social media data analytics to support development. Section 8 concludes the chapter.

2 Web 2.0 Technologies and Social Media

Gartner [12] have defined social media as 'an online environment in which content is created, consumed, promoted, distributed, discovered or shared for purposes that are primarily related to communities and social activities rather than functional task-oriented objectives'. Gartner have also defined Web 2.0 as 'the evolution of the Web from a collection of hyperlinked content pages to a platform for human collaboration and system development and delivery' [12]. Web 2.0 technologies and social media have revolutionalised the way people communicate and socialise. This section provides a summarised discussion of Web 2.0 technologies and social media services.

2.1 Web 2.0 Technologies

Most commonly, Web 2.0 technologies are defined in terms of what they enable people to do, that is, their main agenda, which is to connect people in numerous ways so that they can utilise their collective strengths [7]. In this chapter, the term Web 2.0 technologies is used to refer to web technologies that enable the creation and sharing of social media. The term social media is used to refer to online (Internet) media that is created and shared by communities. Table 1 provides a summary of some of the well-known categories of social media.

As shown in Table 1, different types of social media enable people to express their opinions and get feedback from others (blogs), express their opinions concisely about what is happening right now (microblogs), meet other people and share content with them (social networks), collaborate to generate content (wikis), and upload and share media content (media sharing).

Table 1 A sample of categories of social media

Media category	Purpose	Example of service
Blogs (web-logs)	Facilitate the expression of personal opinions by the public	There are many examples of blogs on the web
Microblogs	Facilitate the expression of personal opinions about what is happening right now	Twitter
Social networks	Professional or social networking sites which facilitate meeting people and sharing content	Facebook, LinkedIn, Twitter
Collaborating	Collaborative reference works (e.g. Wikipedia) that are built using wiki-style software tools	Wikipedia
Media sharing	Facilitate the sharing of digital media, e.g. videos	YouTube

2.2 Social Networks

Dasgupta and Dasgupta [7] have provided a discussion of different types of social networks that existed in 2009. Social contact networks are primarily used for friends and family, e.g. Facebook, Twitter. Study circles are networks dedicated to students. Social networks for specialist groups are used by core field workers like doctors, engineers, members of corporate industries, e.g. LinkedIn. Police and military networks are private social networks (not in the public domain) exclusively for people in these services. There are other network categories which include sporting networks, networks for fine arts, shopping and utility networks. A brief discussion of Twitter, Facebook and LinkedIn social networking services is given in this section.

Twitter: A Social Network and Microblogging Service

A number of microblogging services exist on the web, with Twitter [32] possibly being the most visible (popular) at the moment. Twitter is a 'whats-happening-right-now' social media service [2]. Twitter was launched in 2006 as a microblogging service that allows users to send updates (tweets) to a network of friends (followers) from a variety of devices [20]. Twitter users need a subscription in order to receive updates and tweets are delivered instantly. Tweets are displayed on the user's profile page on Twitter, or they can be delivered via instant messaging (SMS—short message service), Really Simple Syndication (RSS), e-mail or through an application such as Twitterrific or Facebook [20]. Tweets are at most 140 characters. The users interact by *following* updates of people who post interesting tweets. Users can pass along interesting pieces of information to their followers. This is known as *retweeting*. Users can also respond to (or comment on) other people's tweets which is called *mentioning* [6]. The following is an example of a tweet: *RT @toni has a cool #job.* RT is used at the beginning of a tweet to indicate that the message is a *retweet*. Users can reply to (mention) other users by indicating user names prefixed with the @ character (e.g. @toni). Hashtags (#) are used to denote subjects or categories (e.g. #job) [2]. Additionally, emoticons such as smiley (:-), sad face (:-(and variations of these are added to tweets to express sentiment [2, 34]. Tweets may be kept private among the followers, or they may be made public and unrestricted [34].

Twitter is characterised by the large volumes of data that are generated and the large numbers of users. O'Connor et al. [24] have reported that for the 2 year period from 2008 to 2009, the message volume for Twitter increased by a factor of 50. In April 2010, Twitter reported various statistics on the users of Twitter as follows [2, 35] . There were 106 million registered users, 180 million unique visitors every month, and 300,000 new users signing up every day. There were 600 million queries being received daily via Twitter's search engine and three billion requests per day based on the Twitter application programming interface (API). It was also noted that 37% of active users used mobile phones to send requests. More recently, the number of regular Twitter users has been estimated at 200 million.

Facebook: A Social Networking Service

Facebook was launched in 2005 as a social network for use by university and college students in the USA, and so, it was aimed at the young adult age group (18–24 years old). The inventor of Facebook is Mark Zuckerberg, a former student at Harvard University. Facebook opened its services to non-academic users for the first time in 2007. By 2008, Facebook had grown to be the second largest social network with more than 30 million users [15]. As a social networking site, Facebook facilitates meeting people and sharing content such as photos, blogs, microblogs and Facebook applications developed by the users. Big businesses, small businesses, governments and NGOs have all been quick to realise the benefits of having a Facebook presence. Graham [15] has observed that, from a business perspective, Facebook has given big opportunities for businesses to direct their marketing efforts towards the young adult age group (18–24) who are known to be very difficult to win-over.

LinkedIn: A Social Networking Service

LinkedIn is a business-oriented social networking site. LinkedIn users normally associate with their line-of-work network and use the site to maintain a list of contact details for people (connections) they know and trust within their line of work. The network of contacts is used to maintain communication, exchange trade information, academic information, and other types of information. LinkedIn uses a 'gated-access-approach' which means that connecting with other users of LinkedIn requires a pre-existing relationship or the intervention of a mutual contact. This mechanism is designed to create trust among LinkedIn users [25].

YouTube: A Media Sharing Service

YouTube was launched in 2005 and was acquired by Google as a subsidiary in 2006. As stated on the YouTube website (www.youtube.com) the service allows billions of users to discover, watch and share originally created videos. The service provides a forum for people across the globe to connect, inform and inspire others. Video content includes amateur and professional video clips, television clips, music videos, educational videos, and corporate videos. Unregistered users can view videos and registered users can view and also upload videos. According to YouTube Statistics [36], more than one billion unique users visit YouTube every month and millions of new subscriptions happen everyday. YouTube is localised in 61 countries and languages, and 80% of the YouTube usage traffic originates from outside the USA. Mobile devices make up almost 40% of YouTube's global watch time. Individuals, government organisations, small and big business organisations around the world are using YouTube to grow their audiences.

3 Organisations that use Social Media

Organisations have been quick to realise the value of social media. At the present time, business, government and non-governmental organisations typically participate in social media. The reasons for participation include assessing the opinions of the public about products and services that the organisations provide (e.g. Twitter microblogs), soliciting opinions from the public and communication and collaboration between stakeholders (e.g. social networks). This section provides a discussion on how organisations around the world are using social media.

3.1 Big Business Organisations and Social Media Marketing

Big business organisations are routinely using social media for marketing and branding purposes. Stelzner [29] has reported the results of a survey on the usage of social media for marketing purposes. Over 3,000 business organisations worldwide, predominantly in the USA, Canada, UK, Australia and India, participated in the survey. One major conclusion from this survey was that, in these countries, the top five social media platforms for marketing are Facebook, Twitter, LinkedIn, blogging, and YouTube, in that order. According to Stelzner [29], 92% of the organisations used Facebook, 80% used Twitter, 70% used LinkedIn, 58% used blogging, and 56% used YouTube. In terms of usage, Stelzner [29] has observed that (business) organisations conduct the following activities for social media: content creation, analytics, monitoring, obtaining status updates (tweets), research, strategy formulation, and community engagement.

Jansen et al. [20] have conducted research to assess the effectiveness and trends of the use microblogging by businesses for purposes of word-of-mouth branding. They specifically studied the use of Twitter for online word-of-mouth branding. They have concluded that (1) microblogging can be used to provide information to customers and the public in general and to draw potential customers to other online media for the business, such as websites and blogs. So, monitoring microblogging sites concerning a business brand and competitors' brands can provide valuable competitive intelligence information. (2) Using microblog monitoring tools, businesses can track postings and immediately intervene with unsatisfied customers. (3) By setting up corporate accounts on microblogging services, businesses can use microblog polls and surveys to obtain near real-time feedback from customers. (4) Businesses can also obtain valuable product improvement ideas by tracking microblog postings. (5) Businesses can take advantage of contacts made via microblogging services to further their branding efforts by responding to comments made about the company brand.

3.2 Small Businesses and Social Media Marketing

Bodnar [4] has observed that successful small businesses have long thrived on word-of-mouth advertising by satisfied customers in order to help promote their products or services. Due to the ubiquitous nature of the social media services available today, small businesses are in a position to use free tools to help increase word-of-mouth advertising and to decrease the need for expensive advertising channels such as magazines, newspapers, radio and television. Bodnar [4] has reported the results of a study that was conducted on small businesses that have successfully used social media marketing in the USA. Bodnar [4] has also identified a number of challenges and key success factors for small businesses to make effective use of social media marketing. While big businesses have marketing departments that can attend to the time-intensive social media marketing activities, this is not the case for small businesses. However, small businesses can commit weekly resources to creating content and engaging in social media such as Facebook, Twitter or blogs. It is also important for the small business to have some method of establishing how a given social media activity has impacted the business results and to use this information to drive business strategy.

3.3 Public Sector Organisations

Many government departments and other public sector organisations can use social media for purposes of engaging with citizens and conducting polls for various purposes. O'Connor et al. [24] have conducted studies to compare the results of traditional polling with the results of polls conducted via social media, namely Twitter. Using time series analysis, their studies compared the results of public opinion polls (in the USA) on consumer confidence about the US economy with the results of rudimentary (simple) sentiment analysis of Twitter data on these topics over the same time period. They also compared the results of public opinion polls on the popularity of the US president job approval with the results of sentiment analysis of Twitter data on this topic over the same time period. They concluded that a relatively simple sentiment detector based on Twitter data replicates the results obtained using formal and more expensive polling methods. The findings by O'Connor et al. [24] are obviously good news, especially for developing economies. Decision makers should be able to set up polls and surveys, at very low cost, using microblogging services, in order to engage with the country's citizens, assess public sentiment about the economy and other services provided by government departments and agencies. Public safety organisations have also realised the value of microblogging. Jansen et al. [20] have observed that Twitter is increasingly being used by these organisations to receive updates during emergencies and natural disasters so that they can make informed decisions on how to plan rescue operations.

3.4 Non-government Organisations

Non-government Organisations (NGOs) are organisations that are involved in development activities, mostly in developing countries. These organisations consist of a number of stakeholders including: donors, fellow NGOs, staff, local organisations in the developing regions and aid-receiving communities [28]. Sheombar [28] has discussed various opportunities that social media presents for NGOs. These include: collaboration, connecting and interacting, networking, international co-operation, and communicating with the aid-receiving communities via mobile phones. Sheombar [28] has reported the results of a study conducted on Dutch NGOs and their usage of social media and has identified some of the benefits and challenges posed by social media usage. When an NGO has knowledge of the local context, they can specifically target certain groups for purposes of information delivery and data collection. Sheombar [28] has reported that collection of data via mobile phones is a wide spread practice in many development organisations (NGOs). The major benefit here is that organisations are able to reach developing country communities that do not have access to computers. It should be noted that collection of mobile data via mobile phones is not the only data collection method employed by NGOs. Specific crowd funding and fundraising is another activity that has been conducted by NGOs via social media. Some NGOs also continually monitor and analyse their social media activities. Two main challenges identified by the organisations surveyed by Sheombar [28] are the need to respond fast on social networks, and the difficulty of conveying complex messages via social media.

4 Social Media Adoption in African Countries

In the last few years, African countries have experienced a widespread adoption of social media, especially with the young generation. It has been reported in the literature that the key drivers for this adoption have been the widespread adoption of mobile phones, establishment of mobile Internet infrastructure on the continent, and the affordability of mobile Internet services for the ordinary person. This section provides a brief discussion of the drivers for social media adoption and the extent of this adoption.

4.1 Drivers for Social Media Adoption

In August 2012, Deloitte (www.deloitte.co.za) and Frontier Advisory (www.frontieradvisory.com) hosted an African Frontiers Forum to evaluate the economic impact of social media in Africa. They have reported that in 2012, the African continent was the second largest mobile phone market (after Asia) with more than

700 million mobile connections. Estimations of the annual growth were 30% so that by 2016, the number of mobile connections should rise to almost one billion. They have also reported that the widespread adoption of mobile phones and the roll-out of mobile Internet infrastructure in many African countries has resulted in the availability of affordable Internet services to the vast majority of African citizens. This is in stack contrast to Internet access via fixed line telecoms services which are generally unavailable and unaffordable to the vast majority of citizens. Most newly activated mobile devices are Wireless Application Protocol (WAP) enabled. In 2012, Africa's mobile data usage amounted to 14.85% of the total global Internet traffic. It has been argued that the wide adoption of mobile devices, ease of access to Internet services, and affordability of Internet services via mobile devices have been the major driving factors for social media adoption on the African continent [9].

4.2 Social Media Adoption

There has been a widespread adoption of social media in recent years, and Facebook has become the most visited website on the African continent. It was reported in 2012 [9] that for the African continent, the users of the Facebook website were estimated at 44.9 million people. Fuseware and World Wide Worx [11] have reported that in 2014, there are 9.4 million active users of Facebook and 5.5 million users of Twitter in South Africa. It has also been reported that the majority of Facebook and Twitter logins (approximately 80%) from Nigeria and South Africa are from mobile devices [9, 11]. The popularity of Facebook in Africa has prompted Facebook to specifically cater to African markets by starting to roll-out local language versions of the website, starting with Swahili. The Swahili language originates from the East African coast (Kenya and Tanzania) and is spoken widely in East Africa and Central Africa. In addition to the big (American-based) social media services, local social media services have come into existence in some African countries. One such example is Mxit in South Africa (http://get.mxit.com), a social networking and instant messaging service with an estimated user base of six million subscribers [11].

In the private sector, African businesses are increasingly employing social media strategies to engage more effectively with consumers through continuous interaction and engagement [9, 11]. Fuseware and World Wide Worx [11] have reported that, in South Africa, 93% of South African corporations that are major brands use Facebook, 79% use Twitter, 58% use YouTube and 46% use LinkedIn for marketing and branding purposes. However, less than 10% use the home-grown Mxit service. Fuseware and World Wide Worx have further reported that a survey of South Africa's top 50 brands revealed that, on average, they each have 58,000 Facebook fans, 259,000 YouTube account views and 12,785 Twitter followers. In the public sector, many organisations are also using social media to engage with the public. Many government organisations in developing countries have a Facebook

and Twitter presence. As an example, in South Africa, the office of the Presidency [26], the South African government [13, 14] and the National Department of Health [17] all have a presence on Facebook and Twitter. Many public sector and higher education institutions in Africa also have a Facebook and Twitter presence.

5 Online Social Media Analytics Tools

The use of analytics tools is essential for organisations (or individuals) that are serious about building their online presence and growing their organisation's reach online. Many online tools are available for the analysis of social media data. Some tools are specific to a particular service (e.g. Twitter) while others are general purpose tools that can be used to analyse data from different types of social media services. This section briefly discusses some of the available online tools. It should however be noted that online analysis tools tend to come and go rather quickly.

5.1 Online Tools for Analysis of Twitter Data

The Summize tool was a popular online service for searching tweets and keeping up with emerging trends in real time. This tool was acquired by Twitter in 2008 [20]. Summize enabled users to submit queries requesting for the retrieval of tweets on a given topic or brand for a specified time period, followed by analysis of the sentiments expressed in the tweets. Summize would analyse the sentiment and give an overall sentiment rating using a five-point Likert scale with levels (from lowest to highest) wretched, bad, so-so, swell, and great. Twitter Sentiment (now Sentiment140) is an online tool provided by Twitter for online analysis of tweets. Currently (in March 2014), this tool is available at the website http://www.twittersentiment.appspot.com. The tool enables visitors to this site to research and track the sentiment for a brand, product or topic of interest. This website enables a visitor to track queries over time. A visitor can also retrieve sentiment counts over time and retrieve tweets along with their classification. An API is also provided for sentiment analysis [2]. There are also Web tools for searching for tweets. A sample of such tools is given in Table 2. A detailed list is available from [5]. It should be noted again that web tools appear and disappear very often and very quickly.

5.2 Online Tools for Analysis of Different Types of Social Media Data

VentureBeat [33] has provided a brief description of the top ten tools for social media analytics that were in use by organisations at the end of 2013. Three of these

Table 2 Examples of web tools for searching for tweets

Name and URL	What the tool provides
Backtweets, http://backtweets.com	See how many people have tweeted a specific link
NearbyTweets, http://nearbytweets.com	Tweets by Twitter users near you (uses Google Maps)
Twellow, http://twellow.com	Organise tweets in yellow page format, a Twitter directory
Topsy, http://topsy.com	Springboard of more than 30 million users and what they are talking about in real time

tools are briefly discussed here to give an idea of the functionality provided by these tools. Google Analytics is a free resource for social media analytics on an organisation's (or individual's) website. According to VentureBeat [33], Google added *Social Reports* to analytics in 2012. Organisations can use social reports to determine the conversion value of visitors from social sites as well as see how visitors from different social sites behave on the organisation's website. *Social Reports* also has an activity stream that shows in real time how people are talking about the organisation's website on social networks. *Brandwatch* is a tool that monitors all conversations across various social networks. This tool also supports 25 languages. *Hootsuite* is an analytics tool that offers a single online dashboard that an organisation can use to manage their social media accounts such as Twitter, Facebook, Google$^+$, LinkedIn and other accounts. Additionally, tools are provided for social media analytics.

6 Offline Analysis of Social Media Data

Organisations are not limited to the use of online tools for social media analytics. In fact, data can be downloaded from a social network website stored in a special database and then analysed offline using statistical, data mining, and machine learning methods [2]. NoSQL databases are a good candidate for the storage of large quantities of unstructured data such as textual data that is stored by social media services. This section provides a discussion of NoSQL databases and methods that have been reported in the literature for the analysis of social media data.

6.1 Big Data and NoSQL Databases

Social media services have resulted in huge amounts of unstructured data being generated on a continuous basis, as indicated in Section 2. This data is commonly called Big Data. Big data is defined as data with the following characteristics: big volume, big velocity and big variety. Big volume means that the generated data is at scale of terabytes to petabytes. Big velocity means that the data is in motion,

that is, it is arriving at high speed. Big variety means that the generated data is in many forms, that is, structured, semi-structured, unstructured, text, and multimedia data. Web-generated big data is stored in NoSQL databases because relational database systems are not suitable for storing Big Data. NoSQL database systems are distributed non-relational databases designed for large-scale data storage and for massively parallel data processing using a large number of low cost servers in order to provide scalability, availability and fault tolerance [23]. NoSQL databases arose alongside major Internet businesses which had challenges in storing and processing huge quantities of data. Examples of these organisations are Google, Amazon, Facebook and Yahoo! Today they are used by organisations that collect large amounts of unstructured data for analysis purposes. There are currently four categories of NoSQL databases namely: key-value stores, document stores, wide-column stores, and graph databases [23].

Key-value stores store data entries as key-value pairs where the key uniquely identifies the value (data item). The value may be a word, number or complex structure with unique semantics. Document stores (databases) were inspired by Lotus Notes and are designed to store documents. The documents are encoded in a standard data exchange format, e.g. XML, JSON (Java Option Notation), BSON (Binary JSON). The is data stored in key-value pair style but the value column is unstructured data (document). Primary uses of document stores are storing text documents, e-mail messages and XML documents. Two examples of document stores are MongoDB and Apache's CouchDB [23]. Wide-column stores use a distributed, column-oriented data structure which accommodates multiple values per key. These databases use Google's Bigtable structure and file systems (GFS) and MapReduce parallel processing [23]. Graph databases use structured relational graphs of interconnected key-value pairings. A graph is represented as an object-oriented network of nodes (objects), edges (node relationships), and properties (object attributes expressed as key-value pairings). Graph database systems provide visual representation of information as well as an API for querying the graph data. Primary uses of graph databases include representing social networks, generating recommendations and conducting forensic investigations. Examples of graph databases are Neo4j, InfoGrid and AllegroGraph [23].

Organisations that plan to collect and store social media data from social media services should consider investing in NoSQL database systems for storing this data. Document databases (e.g. MongoDB and Apache's CouchDB), as well as graph databases (e.g. Neo4j) are especially suitable for storing social network data. Additionally, MongoDB, Apache's CouchDB and Neo4j are Free/libre/open source software (FLOSS) databases. FLOSS is software that is licensed to grant users the right to use, copy, study, change and improve its design through the availability of its source code. 'Free' refers to the freedom to copy and re-use the software, rather than to the price of the software [10]. Well-known FLOSS projects include Apache web server, GNU Linux, FreeBSD, MySQL, OpenOffice.org and Mozilla. FLOSS offers a number of benefits for organisations. These include reduced software costs, vendor independence and open standards. For developing countries specifically, FLOSS also eliminates the high costs of dollar-based software licences, since

FLOSS licences are much cheaper and are not specific to a machine. There are a number of recognised challenges associated with FLOSS usage. One is the fact that skills are scarce and therefore more expensive. A second challenge is that there is no accountability or possible recourse to legal claims should there be a major problem with the software. A third challenge is the lack of a 24/7 help desk. These challenges indicate that an organisation must weigh the pros and cons of FLOSS before deciding to adopt it for mission critical applications. If the analysis of social network data is not mission critical (and most commonly it will not be) for an organisation, FLOSS database systems should be seriously considered as a viable and affordable solution for data storage.

6.2 Obtaining Data from Social Media Applications

An API is a library of class definitions and functions that enable software developers to access and use the low-level functionality of a given system without having to access the source code. Internet-based companies such as Google, Amazon and Yahoo provide APIs for software developers. Social media service providers such as Twitter, Facebook and LinkedIn also provide APIs that enable developers to develop applications that can access data stored by the service, filter and analyse the data in various ways, and enable other users of the service to use the application. Some online analysis services also provide APIs. Some authors have observed that the use of APIs for web-based services is becoming a trend in application development. Many APIs for web-based application development provide a Representational State Transfer (REST) API. REST is a Web 2.0 standard [7]. REST describes an approach for a client/server architecture which provides a simple communication interface using XML and HTTP. In the REST specification, every resource is identified by a URI and the use of HTTP enables a software developer to communicate through simple GET, PUT and POST commands. This section provides summarised descriptions of the current specifications of the Twitter API and Facebook API. It should be noted that these APIs tend to evolve very quickly.

The Twitter API

Twitter currently provides a streaming API and two discrete REST APIs [2, 31]. Through the streaming API, called the Firehose [2, 22, 31], users can obtain real-time access to tweets in a sampled and filtered form. The API is HTTP based and it supports the use of GET, POST and DELETE requests for data access. In Twitter terminology, individual messages describe the 'status' of a user. Using the streaming API, users can access subsets of public status descriptions in almost real-time including replies and mentions created by public accounts. The streaming API uses basic HTTP for authentication and requires a valid Twitter account. Data can be retrieved in XML or JSON format. The JSON format is very simple and can be parsed very easily because every line terminated by a carriage return contains an

object. The Twitter API allows the integration of Twitter with other web services and applications [20].

The Facebook API

The Facebook API [15] was launched in 2007 and currently consists of five components as follows: (1) an HTML-based markup language called the Facebook Markup Language (FBML), (2) a REST API, (3) SQL-style query language for interacting with Facebook called the Facebook Query language (FQL), (4) a scripting language called Facebook JavaScript, and (5) a set of client programming libraries. The Facebook API enables developers to create external applications to empower Facebook users to interact with one another in new and exciting ways that are invented by the developers. In order to access Facebook data or develop a Facebook application, a developer needs a Facebook account. Obtaining Facebook data for offline analysis is a fairly straight forward matter. One uses the REST API and FQL to obtain Facebook data. Developing a Facebook application however requires a fairly high level of programming expertise. However, Facebook provides many learning aids to help developers to master the use of the API [15]. Since Facebook only provides methods for accessing data and displaying some information to the application user, it is the developer's responsibility to host the application. There are web sites that host Facebook applications for free, although there may be a waiting period after the developer applies for the free hosting [15]. After the application has been developed, hosted at a website and registered (by following the steps to create an application on Facebook and agreeing to the terms of service), Facebook will provide the application to other Facebook users when requested [15].

6.3 Analytics for Social Network Data

Some of the methods of knowledge discovery, data mining that have been applied to social network data are graph mining and sentiment analysis and clustering [2, 34]. This section provides a brief discussion of sentiment analysis, time series analysis and graph mining.

Sentiment Analysis

Sentiment analysis of text messages may be defined as a classification problem where the task is to classify the messages into three categories depending on whether they convey positive, negative or neutral feelings [2, 30]. From a machine learning and data mining perspective, sentiment analysis involves the creation of a classification model which can be used to assign class labels (positive, negative, neutral) to messages. Commonly used algorithms for sentiment analysis are Naïve

Bayes, maximum entropy, support vector machines, and classification trees [2, 30, 34]. Most data mining and statistical software provide these algorithms. Text mining involves the analysis of the message text. The data mining methods that have been used for Twitter text mining include sentiment analysis through classification of tweets, clustering of tweets and trending topic detection [2]. Wakade et al. [34] have discussed the use of sentiment analysis for Twitter data. They have provided a list of activities necessary for sentiment analysis. These activities are given in Table 3.

Activity 1 in Table 3 (data collection) is achieved through the use of the Twitter API to obtain the data about a specific topic. This data may be stored in a document database. Activity 2 requires the use of specialised tools to conduct the data pre-processing activities. The WEKA software [16] provides tools for performing all the tasks listed under Activity 2 [2, 34]. Wakade et al. [34] have observed that two main challenges in the pre-processing of Twitter data are due to the usage of abbreviations (e.g. 'afaik' for 'as far as I know' and 'lol' for 'laugh out loud') and the usage of slang with different dialects such as netspeak and chatspeak. Additionally, it has been observed that in African countries, Twitter users tend to mix English words, French words with words from African languages such as Swahili, Zulu, and many others. One way to address these challenges is to compile additional lists of positive and negative words in African languages. This adds a local context component to the sentiment analysis activities of social network messages. Wakade et al. [34] have proposed the use of an additional pre-processing step to expand well-known abbreviations in tweets. Activity 3 (feature determination) and Activity 4 (sentiment labelling) may require the writing of specialised computer programs to conduct the tasks for the activities. Activity 5 (creation of classification model) can be performed using available data mining software e.g. WEKA [16] and Massive Online Analytics (MOA) [3], or statistical software, e.g. System R [18]. These three applications are FLOSS software and may be downloaded from the Internet.

Usage of the classification model can be conducted using the data stream mining approach. In this approach data (tweets) will be classified as they arrive [2]. The challenge is then to provide effective means of visualisation of classification results by human users. Several writers have observed that the effect of one tweet may be negligible but the effect of many tweets can be significant [19]. The MOA software [2, 3] can be employed for the classification of Twitter data using a classifier created with WEKA and MOA [2].

Time Series Analysis of Aggregate Sentiment

A simple and effective method of providing the results of sentiment analysis classification results to human users in an organisation is through the computation of aggregates (e.g. daily aggregates) which can additionally be displayed graphically, e.g. using line plots or bar charts. Useful aggregates would be the counts or percentages for positive and negative sentiment tweets, or the ratio of positive to negative tweets on a given day. Another even more advanced practice is to use time series analysis. O'Connor et al. [24] have argued that it may be the case that on a day-to-day basis, the aggregate sentiment values may be highly volatile so that it

Table 3 Activities for the creation of a classification model for sentiment analysis of tweets

Activity	Description
1. Data collection	Collect related tweets using a query containing words or phrases denoting the subject (topic) of interest
2. Data pre-processing	Preprocess the tweet data as follows: (1) Remove the stop words. These are words that do not convey any positive or negative meaning. Examples are: is, was, been (2) Replace the URLs with a specific tag. It is common to find URLs in tweets as people share interesting links with friends (3) Remove symbols except those that make up emoticons (4) Stem the words. This is the process of reducing a word to its root form, e.g. the words *reader*, *reading* both reduce to the root word *read* (5) Use the resulting data to create the training dataset from the classification model and a test dataset
3. Feature determination	Determine the features (predictor variables) for the classification model (1) Obtain a list of positive words (e.g. from http://www.inspiration.co.uk/positive.htm) and a list of negative words (e.g. from http://eqi.org/fw_neg.htm) (2) Decide on specific symbols to represent emoticons (3) Combine the positive and negative words as well as the emoticon symbols into the features list (4) Reduce the length of the feature list. Some words in the features list may appear with a very low frequency in the tweets selected for the training data. Use a cut-off value (e.g. frequency > 2) to select the words that appear in the final features list
4. Sentiment labelling	Each tweet in the training data needs to be labelled as positive, negative or neutral. The procedure used by Wakade et al. [34] is as follows: (1) For each tweet, count the number of positive words (*poscount*) and the number of negative words (*negcount*) (2) For each tweet determine whether there is an positive emoticon *hasSmiley*) or a negative emoticon (*hasSad*) or no emoticon (*hasNone*) (3) Use the following logic to label each tweet: Label a tweet as positive if: (*poscount* > *negcount*) *and* (*hasSmiley or hasNone*) or (*poscount* = *negcount*) *and* (*hasSmiley*) Label a tweet as negative if: (*poscount* < *negCount*) *and* (*hasSad or hasNone*) or (*poscount* = *negCount*) *and* (*hasSad*) otherwise label the tweet as neutral
5. Creation of classification model	Using a classification algorithm, e.g. Naïve Bayes, maximum entropy, support vector machines, or classification tree, and the features list, create a predictive classification model and test the model performance using the test dataset

is difficult to determine the general trend for the sentiment measures. They have suggested that the computation of the moving average for the sentiment time series data provides a smoothed measure which makes it easier to observe any emerging trends. The moving average for a time series at time t is computed as

$$MA_t = \frac{1}{k}\left(x_{t-k+1} + x_{t-k+2} + \cdots + x_t\right)$$

where k is the number of past time periods (days in this case) and x_t is the value of the aggregate measure for the sentiment at time t. O'Connor et al. [24] have suggested the use of the ratio (positive to negative) as the aggregate measure for the moving average computations. When the values of the moving average are used for a line plot, the resulting plot is more smooth and more informative than when un-smoothed values are used. O'Connor et al. [24] have argued that smoothing is a critical issue as it causes a sentiment measure to respond more slowly to recent changes, and forces consistent behaviour to appear over long periods of time. It should be noted that too much smoothing (use of very large values of k) makes it impossible to see fine-grained changes to the aggregate sentiment.

Graph Mining

Graph mining is based on the analysis of links between social media users. Data for graph mining may be stored in a graph database (e.g. Neo4j). Graph mining of Twitter data involves the analysis of links between the messages [2]. Bifet and Frank [2] have reported that Twitter graph mining has been used to investigate interesting problems such as measuring user influence and the dynamics of popularity [6], community discovery and community formation in social networks [21, 27], and social information diffusion [8]. These types of analysis can benefit organisations which have a social media presence, to better understand the characteristics of their followers, and to possibly target their most influential followers for purposes of enhancing their online word-of-mouth advertising.

For Twitter data, three measures of user influence that have been reported in the literature are: indegree, retweets and mentions. Indegree influence is the number of followers of a user and directly indicates the size of the audience of that user. Retweet influence is measured through the number of retweets containing a user's name. This measure indicates the ability of a user to generate content which has pass-along value. Mention influence is measured through the number of mentions containing a user's name. This measure indicates the ability of a user to engage other users in a conversation [6]. For Facebook data, the number of fans and the number of posts and comments have been reported as useful measures of influence [11].

7 Discussion and Recommendations

Social media adoption and usage, social networks and data analysis (online and offline) have been discussed in this chapter. This section provides a discussion of the benefits and challenges of using social media, and recommendations on how the challenges can be addressed.

7.1 Benefits That Organisations Can Realise from Social Media Usage

Public sector organisations and NGOs can benefit from the use for social media due to the low cost solutions for providing information to the public, engaging with and obtaining information from the public e.g. through polls and surveys, maintaining a presence on social networks and analysis of social networks data using free online tools. Big business and small business organisations can benefit from the use for social media due to the low cost solutions for marketing, branding and analysis of social media data using free online tools. Public sector organisations and big businesses can further benefit from the use for social media by conducting more sophisticated analysis offline using statistical and data mining tools. Every government has (or should have) a statistician general who heads a department of statisticians. Every big business organisation has a marketing department and public relations department that can dedicate resources to social media usage. However, there are still many challenges that need to be addressed before the benefits discussed above can become a reality for most African organisations. Some of these challenges are discussed below.

7.2 Usage of Social Network Data Analytics in African Organisations

It was stated in Section 4 that many organisations in Africa are actively using social media to engage with the public. FuseWare and World Wide Worx [11] have reported that in South Africa big businesses which are regarded as major brands rely heavily on social media agencies for social media content creation and social media monitoring. However, in 2014, 53% of these businesses plan to build up their social network skills by investing in training for their marketing and public relations teams. FuseWare and World Wide Worx [11] have further reported that the measurement of social media effectiveness by big businesses remains relatively unsophisticated. For Twitter, 83% of these businesses measure effectiveness by the number of followers while only 48% conduct sentiment analysis. For Facebook, 87% of these businesses

measure the number of fans, 79% measure the number of posts and comments, and only 54% are assessing the tone of these posts and comments through sentiment analysis.

7.3 Challenges in Social Media Usage

One major challenge that has been highlighted by researchers is the shortage of expertise in effective social media usage in terms of designing content and conducting social media data analysis to assess the effectiveness of social media usage. This is especially true for African countries and for small businesses. For South Africa, FuseWare and World Wide Worx [11] have reported that, even though 91% of big businesses agree that social media has the potential for building a business, only 19% of these businesses believe that they are getting as much value from social media as they could. This is a strong indication that most organisations (small and large) are still learning how to effectively use social media in order to obtain value from this media. A second challenge that has been highlighted by researchers is the need, for organisations with a social media presence, to continuously monitor the social media services in order to respond to the public in a timely manner. FuseWare and World Wide Worx [11] have reported that, for the top-brand big businesses in South Africa, the average response time for addressing customer issues on Twitter is 271 minutes (4.5 hours). These authors have commented that: 'Taking more than 4 hours to respond to a customer in such an immediate environment shows a gap in social media that needs to be closed'. One obvious conclusion that can be drawn from the foregoing discussion is that, if the big African businesses are still struggling to monitor and respond to social media communications in a timely manner, then small businesses and public sector organisations will also struggle with this aspect of social media usage.

Sections 5 and 6 provided a discussion of methods for obtaining social media data, storing this data in special databases, e.g. NoSQL databases, and conducting offline analysis of this data. At the present time, many organisations use relational (SQL) database, and these organisations possess IT expertise to conduct analysis of data in these databases. Technical expertise in the use of NoSQL databases is still very scarce in most African countries. This is partly due to the fact that university curricula on database systems largely concentrate on the relational database. Expertise in the use of social media APIs, e.g. the Twitter API and Facebook API is also very scarce. Additionally, the functionality of these APIs is changed very frequently by the social media services. In order to build capacity and expertise in conducting the types of social media data analysis discussed in Section 6, it will be necessary for universities and organisations to create educational programs that address these application development skills.

7.4 Research Perspective on Social Media Adoption and Data Analysis

Academic literature on social media adoption and effective usage is hard to come by. It is a worthwhile challenge for African researchers in Computer Science, Information Systems and Statistics to engage in research which has the potential to create a knowledge base on effective social media usage in organisations. Effective sentiment analysis requires the use of sophisticated text mining methods which require (1) lists of positive and negative words (2) pre-processing of text to produce stemmed words, and (3) processing ability to determine the language for words that appear in mixed language (African and European) communications. It is a worthwhile challenge for African researchers in ICT-related fields to engage in inter-disciplinary research with academics in African languages (linguistics) in order to design and implement methods and tools for effectively processing African language text for purposes of text mining. The domain of Web 2.0 computing and social media is highly volatile. It is important for researchers to keep abreast of developments in this domain so that relevant research for economic development can continuously drive organisational understanding, increased adoption and effective usage of social media.

8 Conclusions

This chapter has presented a discussion of the widely used social media and their adoption on the African continent, organisations on the African continent that have adopted these social media, how organisations that use social media can benefit from such usage, and the challenges posed by the adoption and usage of social media. Social media like Twitter, Facebook, LinkedIn and YouTube are widely used by many people all over the world, including African countries. Additionally, home-grown social media services such as Mxit are widely used in African countries. The major benefits of social media adoption by individuals, public sector and private sector organisations are the low cost solutions provided for exchanging information and analysing the effect of social media usage. Two major challenges for organisations have been identified in this chapter. The first challenge is the shortage of expertise in performing the tasks required for effective usage of social media. The required tasks include content creation and data analysis. The second challenge is the shortage of resources for continuous monitoring of social media communications from the public and provision of fast responses to these communications. In order to address these challenges, it will be necessary for African organisations and educational institutions to conduct capacity building activities that can lead to a reduction in the shortage of the required expertise. It will also be useful for African researchers to engage in research activities in support of effective and widespread usage of social media in government, non-government and business organisations.

References

1. P. Anderson, What is Web 2.0? Ideas, technologies and implications for education. JISC Technology and Standards Watch (2007)
2. A. Bifet, E. Frank, Sentiment discovery in Twitter streaming data (2010)
3. A. Bifet, G. Holmes, R. Kirkby, B. Pfahringer, MOA: massive online analysis. J. Mach. Learn. Res. **11**, 1601–1604 (2010). http://moa.cs.waikato.ac.nz
4. Z. Bodnar, Generating small business customers with social media marketing: small business case studies (2014), Hubspot http://www.HubSpot.com. Cited 15 Mar 2014
5. C. Burke, Twitter search tools (2011), http://charleneburke.com/2011/02/12-twitter-search-tools/. Cited 28 Apr 2014
6. M. Cha, H. Haddadi, F. Benevenuto, K.P. Gummadi, Measuring user influence in Twitter: the million follower fallacy, in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 2010, pp. 10–17
7. D. Dasgupta, R. Dasgupta, Social networks using Web 2.0. IBM DeveloperWorks (2009), http://www.ibm.com/deveoperWorks/
8. M. De Choudhury, Y.-R. Lin, H. Sundaram, K.S. Candan, L. Xie, A. Kelliher, How does the data sampling strategy impact the discovery of information diffusion in social media? in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 2010, pp. 34–41
9. Deloitte and Frontier Advisory, Social Media in Africa, in *African Frontiers Forum* (Deloitte & Touche, Johannesburg, 2012), http://www.deloitte.com/assets/Dcom-SouthAfrica/Local%20Assets/Documents/TMT_Social_Media.pdf.. Cited 28 Apr 2014
10. Free Software Foundation http://www.fsf.org. Cited 15 July 2014
11. Fuseware and World Wide Worx, South African Media Landscape 2014 – Executive Summary (2014), http://www.worldwideworx.com/wp-content/uploads/2013/10/Exec-Summary-Social-Media-2014.pdf. Cited 28 Apr 2014
12. Gartner, Gartner IT glossary (2013), http://www.gartner.com/it-glossary/. Cited 28 Apr 2014
13. GovernmentZA, Facebook page for the South African Government (2014a), https://www.facebook.com/GovernmentZA. Cited Mar 2014
14. GovernmentZA, Twitter profile for the South African Government (2014b), https://mobile.twitter.com/GovernmentZA. Cited 28 Apr 2014
15. W. Graham, *Facebook API Developers Guide* (Springer, New York, 2008). http://www.apress.com
16. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutenmann, I.H. Witten, The WEKA data mining software: an update. SIGKDD Explor. Newsletter **11**(1), 10–18 (2009)
17. HealthZA, Facebook page for the South African National Department of Health (2014), https://www.facebook.com/HealthZA. Cited 28 Apr 2014
18. K. Hornik, The R FAQ (2014), http://CRAN.R-project.org/doc/FAQ/R-FAQ.html. Cited 28 Apr 2014
19. IBM Global Business Services, Analytics: the real-world use of big data – how innovative enterprises extract value from uncertain data, IBM Institute for Business value, (2012)
20. B.J. Jansen, M. Zhang, K. Sobel, A. Chowdry, Micro-blogging as online word of mouth branding, in *Proceedings of the 27th International Conference Extended Abstracts on Human factors in Computing Systems,* Boston, Apr 2009, pp. 3856–3864
21. A. Java, X. Song, T. Finin, B. Tseng, Why we twitter: understanding microblogging usage and communities. in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis,* 2007, pp. 56–65
22. J. Kalucki. Twitter streaming API. http://apiwiki.twitter.com/Streaming-APIDocumentation. Cited 15 July 2014
23. A.B.M. Moniruzzaman, S.A. Hossain, NoSQL database: new era of databases for big data analytics – classification, characteristics and comparison. Int. J. Database Theor. Appl. **6**(4), 2013 (2013)

24. B. O'Connor, R. Balasubramanyan, B.R. Routledge, N.A. Smith, From tweets to polls: linking text sentiment to public opinion time series, in *Proceedings of the AAAI Conference on Weblogs and Social Media*, Washington, May 2010, pp. 122–129
25. Z. Papacharissi, The virtual geographies of social networks: a comparative analysis of Facebook, LinkedIn and ASmallWorld. New Media Soc. **11**(1/2), 199–220 (2009)
26. PresidencyZA, Facebook page for the Presidency of the Republic of South African (2014), https://www.facebook.com/PresidencyZA. Cited Mar 2014
27. D.M. Romero, J. Kleinberg, The directed closure process in hybrid social information networks, with analysis of link formation on Twitter, in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 2010, pp. 138–145
28. A. Sheombar, Social media use by Dutch development organisations: perceptions, applications and barriers. in *CDI Seminar,* Nov 2013
29. M.A. Stelzner, *2013 Social Media Marketing Report: How Marketers are Using Social Media to Grow Their Business*. Social Media Examiner, SocialmediaExaminer.com, (2013)
30. S.L. Ting, W.H. Ip, H.C. Tsang, Is Naive Bayes a good classifier for document classification? Int. J. Software Eng. Appl. **5**(3), 37–46 (2011)
31. Twitter, Twitter API. (2010), http://apiwiki.twitter.com/
32. Twitter, Twitter website: http://www.twitter.com (2014)
33. VentureBeat, Top 10 social media analytics tools: the VentureBeat index (2013), http://venturebeat.com/2013/12/20/top-10-social-media-analytics-tools-the-venturebeat-index/
34. S. Wakade, C. Shekar, K.J. Liszka, C.-C. Chan, Text mining for sentiment analysis of Twitter data, in *International Conference on Information and Knowledge Engineering (IKE'12)*, 2012, pp. 109–114
35. J. Yarow. Twitter finally reveals all its secret stats. BusinessInsider Weblog Article (2010), http://www.businessinsider.com/twitter-stats-2010-4/
36. YouTube Statistics (2014), http://www.youtube.com/yt/press/statistics.html. Cited 28 Apr 2014

Electronic Commerce in Egypt

Sherif Kamel

Abstract The information and communication technology (ICT) evolution is affecting many nations around the world through one of its emerging transformational platforms, the digital economy. Therefore, electronic commerce (eCommerce) is increasingly becoming the way to trade and do business in the twenty-first century. It reflects a concrete example where emerging technologies can contribute to business, socioeconomic development, and growth for different societies. However, there are multiple challenges that exist that are creating a digital divide between the haves and the have nots that call for more inclusion and more effective mechanisms to deploy the advantages of technology so that the developing world capitalizes on the potentials of the much promising global marketspace. Africa is a continent with massive potential and its economy could be boosted if a state-of-the-art eCommerce platform is in place. Respectively, Africa's emerging countries, including Egypt, with its economy in transition, started investing in building its information, communication, and technology infrastructure since 1985 as a platform for national development. This chapter describes the emergence of eCommerce as a vehicle for business development and the opportunities it represents for Africa with an emphasis on one of its emerging economies, Egypt.

1 Overview

Over the last few decades, the information and communication technology (ICT) became vital as a platform for business and socioeconomic development. Moreover, the Internet became an important medium for information acquisition and knowledge dissemination across the globe [13], which led to the formulation of the global information society and the creation of the digital economy with its growing trends such as competing in time, blended learning, telecommuting, and smart communities among other elements. However, billions of people are still unconnected [31]. Moreover, eCommerce, being the product of the digital economy,

S. Kamel (✉)
School of Business, The American University in Cairo, AUC Avenue, P.O. Box 74, New Cairo 11835, Egypt
e-mail: skamel@aucegypt.edu

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_7

remains more like a promise than a reality at least within the context of developing countries [30], with so many opportunities are still untapped that relate to trade or to the larger context of the digital economy including electronic business (eBusiness) and electronic government (eGovernment) and others.

Decision and policy makers, scholars and practitioners for many years now, have acknowledged and emphasized the important role eCommerce can play in economic and societal development especially in the context of developing and emerging economies [1]. It is important to highlight the role of the government in setting-up the required and enabling legal framework that can enable eCommerce. Moreover, the emphasis on human capital capacity development and the formulation of effective financial mechanisms to enable eCommerce is best realized through collaborative work between the government, the private sector, the civil society, and other stakeholders in the community.

eCommerce helps people do business easily and quickly [3, 27]. The immediate nature of communication through computers, notebooks, tablets, and mobile phones reduces the cost of transactions because of the virtual non-paper-based environment while providing an extended set of options for different customers that is vendor-free and location free. eCommerce allows business and entrepreneurs to become more competitive locally and globally and can gradually and positively transform people's lives. In addition, there are opportunities for eCommerce to contribute to the development process and to assist in narrowing the digital divide and the gap that was created between the haves and the have nots [9]. Based on the developments that took place across the world over the last few decades in terms of ICT deployment, one of the regions that could be the venue for future growth and development in the twenty-first century is the African continent.

2 Africa Going Digital

Africa is more than a one billion-people continent with massive growth potential. It is arguably the continent of the twenty-first century. It is worth noting that the development in digitizing Africa has seen massive improvements over the last few decades. However, there are still millions in Africa who are not connected to the Internet and respectively they are not benefiting from the opportunities enabled by the global marketspace. Moreover, given the multiple challenges including the landline constraints and unaffordable prices of personal computers or notebooks, many users in developing economies such as the case in Africa resort to Internet cafés and the use of mobile phones for eCommerce browsing and transactions [5]. Electronic payment (ePayment) remains an issue despite the formulation of a number of solutions but overwhelmingly eCommerce in developing countries rely primarily on the provision of information rather than online transactions. In many cases, orders are made online and payments are made upon delivery offline. The blended model is widely diffused for multiple reasons including lack of complete

infrastructure, trust, and security issues. The legal and regulatory framework remains a necessity and represents a major challenge to the diffusion of the digital economy. Moreover, it is important to note that one of the primary technology infrastructure that is necessary for eCommerce to grow and disseminate is the provision of an ePayment mechanism that is universally accessible and available. Today, due to the lack of such infrastructure, coupled with the relatively small credit card community, no less than 80 % of the transactions are based on cash on delivery. Therefore, for eCommerce to grow exponentially, ePayment needs to be diffused across the community.

eCommerce has the potential to transform the economic landscape of different societies. However, countries that will fail to adopt the concepts of eCommerce run the risk of losing out on multiple opportunities that can support their socioeconomic development plans. There are many issues that remain unsolved, infrastructure is not complete, and readiness levels are below par as compared to other countries outside the African continent. In addition, understanding eCommerce within developing nations remains a challenge to deal with [28]. When comparing the developed versus the developing worlds; while the first have managed to harness the potential and opportunities enabled by ICT in general and the Internet in specific through eCommerce, the latter is lagging behind in terms of building its national information infrastructure, engaging in the digital economy and focusing on providing low-labor cost to the developed world [33].

In 2003, 95 % of Internet and eCommerce technologies were adopted by the leading developing countries [34]. The key is always people, their awareness, training, and willingness to accept technology and capitalize on the offerings it provides. For example, countries like Estonia, an emerging economy like many in Africa, have technology acceptance and usage average rates that is equal to many developed economies while India, despite its numerous technology professionals, has much less acceptance and usage rates [20]. The element of social influence also becomes important in the mix when it comes to technology acceptance. In that respect, eCommerce could be perceived in terms of its adoption, diffusion, and adaptation elements and how that will bring about social development [5]. It is important to note that the deployment and diffusion of telecommunication facilities and infrastructure in remote and underprivileged communities to get access to different ICT tools and applications can empower people and provide alternatives for socioeconomic development [12, 22]. Moreover, eCommerce in developing communities opens new opportunities for the local people to engage and contribute to their societies and tap into other development vehicles that were not available before. Technology represents an opportunity for younger and relatively smaller enterprises where different tools and applications help bridge the competitive advantage larger and more established businesses enjoy.

Social influence plays an important role in understanding the adoption behavior in different societies with an emphasis on developing economies in their quest to close the gap with developed economies [32]. Consequently, according to Datta [5]: “social influence is defined as the degree to which individuals in developing nations

perceive how other people they consider important feel about their adoption of eCommerce technologies." In this context, there is a handful of developing nations that would tend to adopt information technology (IT) to raise its status among countries with similar socioeconomic conditions in the developing world [23].

Moreover, in terms of policy-related issues, the overall business, industry, and economic landscape should enable eCommerce to grow and prosper. Within the context of developing economies, the level of competition, privatization, size and magnitude of the private sector, the promotion of youth, start-ups, innovation, entrepreneurship, the use of ICT, and the legislative framework are all important aspects that have concrete and effective implications on the deployment and growth of eCommerce [5]. The degree of education and level of awareness across the community is always an important determining factor in the successful diffusion and adoption of eCommerce. It is important to note that many African countries have taken real steps in building their infrastructure since the late 1990s when most developing countries including the ones in Africa were classified as having poor telecommunications, poor transport system, poor ePayment systems, and minimal skilled workforce [24]. Investing in human capital is key. The process starts at the school level and continues throughout university education. At this time, ICT is not an option. The need to provide continuous awareness, training, education, and lifelong learning is extremely important and should be embedded throughout the curriculum of different schools, universities, and different educational institutions.

3 Electronic Commerce in Africa

In recent years, there have been multiple developments in eCommerce in Africa. The question remains critical, is Africa ready for eCommerce? Is the right investment in place? Are people engaging with unconventional ways to trade and do business? Increasingly with the notion of globalization and internetworking, individuals around the world can do business and trade anywhere and anytime through the digital space which represents a huge opportunity for African economies. Unfortunately, despite the improvements taking place in different parts of Africa, the continent statistically remains high on the poverty indices and in terms of Africa's stake in the space of global trade that too remains negligible compared to other regions in the world. Therefore, the findings are not promising yet when it comes to assessing the possible immediate benefits that could materialize through eCommerce. It is important to note that many of the business transactions remain paper-based, involving cash payments and usually face-to-face meetings between different partners and that applies to different sectors and industries [10].

There is no doubt that the Internet connectivity evolution that is growing across Africa is yielding an increasingly huge consumer base that represents a massive potential for Africa's digital economy. However, there is still a lot of work that needs to be done first because today out of Africa's one billion population, less than 20 % are connected to the Internet. The challenges include completing the infrastructure,

the legal framework, payment methods, and other core services that need to be offered. The future will probably relate to mobile commerce (mCommerce) with an increasing number of Africans turning to their phones and other mobile devices for various services and would probably be trading and purchasing their needs using their mobiles if the infrastructure is available. The growth of smart phones and the diffusion of tablets is also helping in that direction. In the future, it is probably mCommerce that will drive the digital commerce interactions with the formulation of hybrid models that would serve the different conditions across the African continent. This will also help in creating the culture that will pave the way for as many Africans as possible to migrate from the marketplace to the marketspace.

For eCommerce to make a difference there is a need for an adequate framework that caters for the marketspace [6]. The key remains in the legal setting and the formulation of laws and policies that would promote and encourage the development of a vibrant space for eCommerce to grow and flourish [10]. According to the Organization for Economic Co-operation and Development projections, Africa's contribution to global trade was about 5 % in 2011 [25]. During the last two decades, there has been increasing appreciation and awareness to the potentials that ICT can create through emerging platforms, tools, and applications such as eCommerce and its role in transforming economic development and growth in the African continent. Every year, an increasing number of Africans go online, navigate the cyberspace, actively contributing in the social networking world and engage digitally. Investments is growing across different stakeholders including the providers and the users. Such growth should be sustained for a number of years for eCommerce to have the scalable and sustainable impact desired.

In the case of South Africa, there has been a boom for a number of years. In that respect, there is a blend of traditional brick and mortar as well as online companies including a variety of online marketplaces, specialty stores and a number of outlets in sports, auctions, food, and the traveling and tourism industry. It is important to realize that eCommerce is fast gaining ground as the accepted and convenient business platform. There are also a number of virtual shopping malls that are increasingly having a growing role in African markets. eCommerce as an ecosystem has the ability to empower young and skilled African entrepreneurs to venture into innovative ICT-based start-ups that can help transform African economies. With the removal of time and distance barriers using ICT, the world becomes a marketspace that is open for solutions that can meet the growing and diversified consumer demands. This could include trading, software development, consulting, financial services, entertainment, education, health services, and many more [10]. The world is increasingly living in a space where access to information provides you with an opportunity to shop, decide, buy, and pay online 24/7 from anywhere and through different means [7].

One of the primary issues that needs to be addressed when building Africa's infrastructure for eCommerce is the financial building block. There is a need for a much more advanced and enabling environment. Online payment processing

is a critical success factor for the formulation of a successful and sustainable eCommerce set-up. The issues of Internet connectivity, percentage of usage per capita, and bandwidth are all on their way to be resolved due to the efforts exerted by both the governments and the private sector in the continent, but the online payment remains a major challenge. There are some success cases that could be used as role models such as Kenya, one of the most advanced countries in Africa when it comes to eCommerce with their M-PESA platform. In addition, Rwanda, given the commitment of its government and private sector companies in promoting eCommerce, they are putting pressure on banks and financial institutions to create mechanisms that can help develop and grow online transactions that could pave the way for a digital economy.

In the case of Ghana, with its low level of development, the country is improving gradually in terms of ICT infrastructure with multiple projects that address different sectors including renewable energy as well as virtual shopping malls within rural communities. There is also a successful roll-out of MTN mobile money in Ghada and Uganda. In addition, Zambia launched Africa's first mobile payment service in 2002 and its success led VISA International to acquire it in 2011. In the case of Egypt, despite the recent economic and political developments, innovation is still popping-up in different forms such as the mobile wallet service that was launched in 2013 by MasterCard, the National Bank of Egypt, and the country's mobile operator Etisalat. The application represents a breakthrough in a nation with 94 million mobile subscribers' lines served with an innovative interoperable Arabic mobile money application that allows subscribers to pay their bills through their mobile phones, make eCommerce payments, top-up their mobile prepaid lines, and pay for different goods and services at several shops across Egypt.

Moreover, Africa can transform its travel and tourism industry through a blend of mCommerce and eCommerce applications offering a variety of tools and payments methods. It is important to understand that Africa is not a nation but rather a continent with differences in preferences, cultures, and values across its many countries. In terms of payment, Africa is characterized by huge diversity in payment methods and references which clearly indicates the need for adaptation and that there is no one size that fits all. For example, while M-Pesa is popular in Kenya where there are much more mobile money accounts than bank accounts, the experience was totally different in both Nigeria and South Africa. The experience in Nigeria is Verve cards which clearly outnumber credit cards such as Master Card and Visa. Such diversity and the associated vulnerability and risk reflects the fact that to do eCommerce in Africa, companies and individuals need to be very careful and adapt different strategies and approaches according to the conditions of each and every country.

eCommerce can also be a platform for rural development and poverty alleviation across the African continent creating opportunities for underprivileged communities allowing the formulation of entrepreneurial activities for trading for farmers and other groups. eCommerce and especially mCommerce could provide an ideal diffusion platform for promoting different products and services on offer. In the

post January 2011 period, this is being helped by the proliferation of ICT and Internet-based start-ups offering a variety of services and products that addressed different societal and market needs and used by a growing technology-savvy young population in Egypt.

4 Building the ICT-Enabling Ecosystem in Egypt

Since 1985, Egypt has invested in its ICT infrastructure targeting the build-up of its national information backbone to become the platform for the development of all economic sectors based on timely, relevant, and accurate information. During the period 1985–1995, a public–private sector partnership helped realize the establishment of Egypt's information highway [13]. The program embedded the establishment of hundreds of informatics projects and centers in different government, public, and private sector organizations as well as the development and improvement of all the building blocks of the information infrastructure such as people, technology (hardware and software), networks, information, and knowledge management aspects [15].

In 1991, the government began an economic reform program aiming at transforming the Egyptian economy from centrally planned, inward looking economy to one that is market-based and internationally oriented [18]. It paid dividends and during the late 1990s and the first decade of the twenty-first century, economic growth was multiplied several times reaching 6 %. The government's policy had focused on the removal of price distortions and obstacles to investment and trade and worked on a plan to introduce smooth and effective processes for the reformation of the financial sector. The important features of the program were the deregulation of foreign exchange, budget deficit financing, the gradual removal of government subsidies to cut down on expenditures, the implementation of a privatization program, the introduction of a capital market law, the abolishment of investment licensing, and the revision of the trade policy through the reduction of the level of tariffs of the GATT [18]. The focus of the government of Egypt was in the development and realization of the digital economy. At the time, this had taken a further boost with the appointment of a cabinet in 1999 that was geared towards investment in the development of an economy that capitalizes on the benefits of ICT. However, post January 2011 and Egypt's uprising, the economy clearly slowed down. The promising sign is that one of the very few sectors that were not really affected by the developments taking place in Egypt was the ICT sector. The following table demonstrates the statistics of the ICT sector.

The Internet is a major driving force of change in the global market place [13]. It is a global information highway linking countries such as Egypt with the rest of the world [2]. It changed people's lives in the way they work, study, shop, and get entertained among other diverse implications. The growth of the number of Internet users has been stunning over the last three decades and is expected to continue to rise as the world becomes more aware of the opportunities enabled by the Internet.

Indicators	Oct 1999	Dec 2004	Dec 2006	Dec 2008	Dec 2011	Dec 2012	Dec 2013
Internet subscribers	300 K	3.6 m	6 m	11.4 m	29.8 m	36.8 m	37.1 m
ADSL subscribers	N/A	N/A	206 K	593 K	1.65 m	2.24 m	2.49 m
Internet penetration per 100 inhabitants (%)	0.38	5.57	8.25	16	35	44.11	44.2
Mobile phones	654 K	7.6 m	18 m	38 m	79 m	97 m	97.5 m
Mobile phones penetration per 100 inhabitants (%)	0.83	9.74	23	51	98	117	118
Indicators	Oct 1999	Dec 2004	Dec 2006	Dec 2008	Dec 2011	Dec 2012	Dec 2013
Fixed lines (m)	4.9	9.5	10.8	11.4	8.96	8.56	6.84
Fixed lines penetration per 100 inhabitants (%)	6.2	12.1	13.8	15.2	12	10.51	8.28
Public pay phones (K)	13	52	56	58	24	15	14
IT clubs	30	1,055	1,442	1,751	2,163	2,163	2,163
ICT companies	870	1,870	2,211	2,621	4,250	5,083	5,237
IT companies	266	1,374	1,970	2,012	3,599	4,116	4,245
Communications companies	59	152	244	265	295	375	390
Services companies	88	148	211	242	356	592	602
Number of employees in the ICT sector[a] (K)	48	116	148	174	212	216	221

Electronic Readiness in Egypt (www.egyptictindicators.gov.eg) [8]
[a]There are over 60,000 indirect workers in both IT clubs and Internet cafés

In 2013, there were 200 million Internet users who went online for the first time. Today, in Egypt, over 45 % of Internet users connect via mobile devices [29]; this compares to 36 % in 2011 and 28 % in 2009 [21]. By the end of 2011, mobile Internet users represented 13 % compared to 8.9 % in 2009 [21]. In addition, broadband is spreading universally among rural and urban areas with Cairo and the main cities still having the largest density of users [21].

Thus, Egypt has been striving to implement a long-term strategy to support the realization of its targeted development objectives using state-of-the-art ICT [15]. Today, the Internet penetration in Egypt stands at around 44 % [4]. The associated cost, which directly affects the growth of eCommerce, is relatively expensive where an individual would pay for a speed of 2 MB/s for unlimited data US$32 per month (US$384 per year). It is important to note that the per capita income in 2010 in Egypt was about US$3,600. The government of Egypt, based on its set strategic plan, hopes to extend Internet penetration to 75 % of households by 2015. It is important to note that post-2011 in Egypt, social networking is emerging as a widely popular activity. In December 2010, there was 4 million Facebook accounts in Egypt reaching 11 million accounts in July 2012, with 75 % of those under the age of 30 [29]. There is also an increasing number of Twitter and Google + users, all contributing to a community that is becoming more Internet-oriented and information-driven, especially among the younger generations. The evolution of social networking is also a result of the increasing diffusion

of smart phones in Egypt since 2012. The social networking phenomenal growth, from 6 to 20 million in 18 months, allowed social inclusion through the digital world for different groups including women, youth, and underprivileged communities in remote areas in Egypt. Since 2008, Egypt's youth realized the power of sharing information and using emerging ICT tools such as social networking [17].

Historically, the Internet started in Egypt 1993 with 2000 users [14]. As an attempt to diffuse the Internet usage among the society, the Cabinet of Egypt Information and Decision Support Center in collaboration with the Regional Information Technology and Software Engineering Center provided free Internet access on a trial basis to the public, private, government, and non-government organizations to entice the users to venture into the new technology. This was done with the financial support of the government, in an attempt to aid in the global exposure of the market and to pave the way for the commercialization of the Internet services. The free access formula was accredited for contributing to the boost in the rate of growth of Internet users, especially within small and medium sized enterprises (SMEs) and industry and sector professionals. An early attempt to promote entrepreneurship and a start-up society. In 1996, the government replaced its free Internet access policy with an open access policy and Internet services for the commercial domain were privatized, and 12 Internet service providers started their operation.

The twenty-first century promises to bring to the world more innovations, more opportunities and more challenges. Therefore, countries around the world, including Egypt, should be prepared for a more competitive global marketplace that is information-driven. This is an invaluable opportunity for the developing world. However, the information society in general represents both a challenge and an opportunity for the developing world. It has the potential to be socially beneficial in issues related to economic growth, education, and business development and can also help in alleviating poverty, improving access to healthcare, fairly distributing resources, and strengthening participation in decision-making processes. Therefore, the impact of the Internet should be measured less in terms of sheer numbers of connected individuals and more in terms of value-added contribution to social progress. The movers and shakers will be those who can capitalize on innovative ICT in growing business and trade. Historically, this development compares to what occurred over a hundred years ago, when the world's economy evolved from an agricultural society to an industrial society; today, the world is moving into the information society and countries need to be ready, Egypt included.

5 Electronic Commerce in Egypt

eCommerce, like the Internet, brings about the same type of decision, where it provides unprecedented opportunities for increasing trade, promoting investment, facilitating business transactions, providing a larger and more varied market, and supplying an unparalleled marketing tool for developing economies. In fact, eCommerce represents an opportunity for the developing world to close the gap with

the developed world given the nature of marketspace provided that the ecosystem required is in place. It also provides an opportunity for SMEs to compete at a wider scale while capitalizing on emerging ICTs. In that respect, eCommerce carries strong business and socioeconomic implications for Egypt and provides many opportunities to access global markets. Moreover, with 58 % of the population under the age of 25 where a large majority is entrepreneurial, technology savvy, and interested in start-ups and deploy ICT tools and applications to address different socioeconomic issues, the future looks positive. In Egypt, since 1998, there were many click and mortar companies established providing virtual commercial malls and online shopping sites [26].

In the early days of eCommerce diffusion in Egypt, there were a number of examples such as www.masrawy.com,[1] which is a portal with contact information of different businesses in Egypt coupled with updated information on the stock market and other news updates. There was also www.otlob.com,[2] which is the first online food delivery service covering over multiple restaurants in greater Cairo. Otlob.com is a perfect example of Egypt's quickly emerging eCommerce community from a business2consumer (B2C) perspective. The service is fast, convenient, and was launched in 1999 and is still growing and managed to diversify its services and offerings while expanding across Egypt geographically and adding multiple services [19]. Today, Otlob services include information on more than 200 restaurants and the operation represents around 7 % of the total food delivery service in Egypt [29]. There are other examples in Egypt that are all directed and channeled towards the development of electronic communities. Currently, there are many sites that offer a blend of online and offline services but way below the potential of the country given its demographics demonstrating the early stage of eCommerce Egypt still have. eCommerce represents a dynamic model for the information society and an opportunity for Egyptian firms to increase competition, which will be reflected in improving the level of quality; create opportunities for employment; enforce competition; facilitate participation at the international level; increase government transparency; and diffuse the use of IT.

eCommerce could enable Egypt to experience a more open economy and increase its comparative advantage worldwide; hence, helping in the nation's economic development, and providing new opportunities for penetrating international trading markets especially for SMEs, which lack the resources enabling them to promote itself globally [16]. Since the web's commercial growth in 1995, the growing popularity of eCommerce in Egypt has changed the way computer users shop for goods and services [11]. However, the volume has not even come close to the average potential in a nation of more than 82 million people. Therefore, in its strive to lift-up its developmental process, Egypt has formulated a national plan demonstrating its vision to prepare itself for a more competitive and global market environment that is enabled by the information age through investing in the

[1]Masrawy means Egyptian (in Arabic).

[2]Oltob means order (in Arabic).

build-up of a comprehensive eCommerce infrastructure. The long-term objectives of the plan aim at formulating a framework for eCommerce coherent with national policies and compatible with the global market, developing the national information infrastructure required, improving awareness, and stimulating cooperation between the government, the industry, and the private sector and promoting online business development platforms.

To promote eCommerce, Egypt adopted multiple strategies to increase the diffusion of the Internet leading to an average growth between 2000 and 2011 of 3.2 % reflecting the Internet demand side [21]. From an Internet supply perspective, the international bandwidth increased with an average annual growth rate of 97.3 % during the period 2001–2011 to reach 181,854 MBPS yielding a per capita share of Internet bandwidth of 2249.22 MBPS [21]. The same period witnessed some dramatic increase in broadband Internet both in urban and rural areas. Internet usage is regularly increasing among government agencies and corporates. In 2011, less than 2 % of Internet users shopped online due to lack of understanding, familiarity and also because of online trust and security reasons [29]. Most recently in March 2011, Offerna became the first group-buying website in Egypt, following the Groupon concept, adapting its services to the local market. The company enjoyed 40 % annual growth rate in revenues since then and positively contributed to building trust in the eCommerce space in Egypt [29]. One of the lessons learned from the experience from Offerna in Egypt was capitalizing on social networking which gained dramatic momentum through Facebook and Twitter.

Moreover, the digital economy is being demonstrated through having more than 82 % of the main government institutions using the Internet in 2011 compared to 76 % in 2010, achieving an average annual growth rate of 6.6 % [21]. It is important to note that almost 25 % of the main government institutions using the Internet have participated and/or offered eGovernment services in 2011 compared to about 9 % who participated in eCommerce activities [21]. The eCommerce services undertaken by the main government institutions included, but not limited to, receiving orders of selling goods/offering services; sending orders of selling goods/offering services; advertising goods and services; bids and tenders publication and receiving technical and financial offers. With respect to the private sector, Internet usage has been increasing proportionally with firm size. For example, on the one hand, 86 % of the relatively larger companies (250+ employees) are using the Internet which is the highest among the private sector companies in Egypt; on the other hand, 31 % of small enterprises (10–49 + employees) are using the Internet [21]. Moreover, broadband is dominating private enterprises Internet access (around 91 % of private business enterprises).

eCommerce promises to offer enormous opportunities for the market in Egypt with implications on the labor market where new employment potential will focus on information-based and value-added services that can have a vital impact on business and industry growth. There is a growing interest among the business community shown by moving to Internet-to-market and promote ideas, products, and services. There are great opportunities for eCommerce dissemination in Egypt and it will change the way business is done whether business2business (B2B) or

B2C such as in the banking and financial sector where recently banks are showing a growing interest in online transactions and in supporting online businesses but it is still in the early phases.

In terms of electronic sales (eSales) and electronic purchases (ePurchases), they represent 30 and 16 % respectively of private business enterprises total sales [21]. In addition, 47 % of the private businesses enterprises using the Internet provide customer services online while around 32 % of them use the Internet for online banking and financial services.

6 The Digital Economy in Egypt

The digital economy in Egypt is still in its early days. However, although small, there is a dynamic eCommerce platform in Egypt. In 2011, the Internet contributed to the economy around US$2.2 billion that was comparable to 1.1 % of the nation's gross domestic product (GDP) [29]. The positive implications is that the development of the digital economy was fueled by entrepreneurship, innovation, creativity, investment promotion, and technology start-ups, and supported by different stakeholders including the government, the private sector, and the civil society. There is a huge potential, especially with the proliferation of angel investment and venture capital networks.

The ICT sector in Egypt has generated revenues of around US$9.3 billion during fiscal year 2011–2012 compared to US$8.7 billion during fiscal year 2009–2010; during that period, ICT revenues growth rate was calculated to be 6.5 % [21]. It is worth noting that the ICT sector makes a substantial contribution to Egypt's GDP with 4.2 and 4.6 % during fiscal year 2009–2010 and 2011–2012 respectively [21]. The ICT sector achieved the highest growth rate in 2010–2011 (7 %) and the second highest in 2009–2010 (13 %) among different socioeconomic sectors in Egypt [21]. This has also been reflected in the number of ICT companies established and the job opportunities created in the sector.

While today the consumer's total digital purchases make around 50 % of the Internet contribution to the economy in Egypt, projections indicate that the total contribution will grow three folds by 2017 to reach 1.6 % of GDP [29]. However, it is important to note that Egypt's businesses and industries have not yet grasped the significant commercial potential and capacities of the Internet. There are sectors that could benefit clearly from the digital economy such as the travel and tourism sector where the industry's potential online marketplace for travels and tourism services of US$1.8 billion of which Egyptian tourism companies exploit less than 5 % [29].

eCommerce is estimated to contribute US$229 million representing 0.1 % of Egypt's GDP [29]. This modest volume represents a major divide with other countries that are gradually developing their Internet economy. For example, in 2011, online advertising was estimated to be 4 % of Egypt's ad spending [29]. This demonstrates a model of B2B economic activity that the Internet generates and increasingly challenges the traditional TV ads space which represents the nation's

most popular and widespread advertising platform. In addition, the digital economy is an opportunity to leverage social services and promote social inclusion. For example, the government of Egypt used the Internet to improve the efficiency in delivering services such as food subsidies through the "family card" to monitor the purchase, transparency, and fairness of food distribution [29].

There is no doubt that countries in Africa can work together and learn from each other. With the diversified culture across different countries, more added value could be brought into a blended model that could see a market that is populated with more than one billion people benefit from a continent-based marketspace for the African region through B2B; B2C and business2government (B2G) transactions. While the associated investment is doable through private–public sector collaboration with the support of the civil society, the potential return is massive making the return on investment (ROI) all worth it and encouraging all stakeholders to contribute in making the digital economy a reality for the emerging continent, Africa, the continent of the twenty-first century.

7 eCommerce Challenges in Egypt

For eCommerce to realize its targeted business and socioeconomic development objectives, there are a number of challenges that need to be faced that relate to a variety of social, technical, financial, and legal elements. With respect to the social challenges, there is lack of awareness and training, lack of trust, resistance to change, and the language barrier. Awareness is considered a major deterrent and that includes customer and organizational awareness of the benefits of the Internet and the digital economy. The lack of training is also a major obstacle where people are not prepared to handle operations in the marketspace, which creates confusion for those accustomed to traditional settings when they are introduced to innovative techniques. The lack of trust remains another challenge with respect to ePayment given the cash-based society in Egypt. Resistance to change is another factor that has more of a cultural aspect where people find it difficult to change their habits. Finally, there is the language barrier where the online availability of Arabic content remains low [29].

With respect to the technological challenges, there is still the problem of relatively weak resources with respect to the telecommunications infrastructure. This includes bandwidth cost that is almost two and half times more expensive than the international tariff with low capacity level, which is extremely modest leading to long time for access and downloading. With respect to the financial challenges, there is the lack of ePayment systems. For eCommerce to succeed, ePayment should be available, efficient, and secured. In addition, there is also major deterrent represented in the low penetration rate of credit cards in Egypt creating a major challenge since credit cards are the primary method of setting consumer transactions over the Internet. Customs and taxation represent another barrier for eCommerce diffusion. With respect to the legal challenges, there is a need for the

development of a legal framework for eCommerce operations in Egypt since the Egyptian law did not yet adapt to the norm of environment of trade transactions in the digital economy. It is important to note that the above-mentioned challenges need to be dealt with strategically at the national level and through proper allocation of resources, many of which need to be transformed into opportunities that will then enable the development of a solid infrastructure of eCommerce in Egypt.

8 Related Work

The challenges that face eCommerce in Egypt are virtually the same in Africa and in most of the developing world. There is no one size that fits all but there are definitely lessons that could be shared. In different countries and regions around the world, eCommerce, eBusiness, and eGovernment have become platforms enticing SMEs and other forms of companies and NGOs to compete with larger organizations because of the advantages the marketspace provides for everyone. The enablers and regulatory frameworks have been key in moving from the marketplace to the marketspace and encouraging a larger segment of the community to trade and shop online. The margins online are relatively smaller and provide a unique opportunity to start-ups and smaller companies to compete at the highest level with the big brands and multinational companies. In many communities in Egypt and other African nations, the cell phone itself is becoming a platform around which businesses are being created without even physical locations; it is all about communicating, human capacities and skills, and the proliferation of businesses through mCommerce.

The future will be determined not by the most state-of-the-art technology used but rather by the effective use and the impact such use creates across the community. eCommerce benefits from a global marketspace irrespective of the time and distance barriers and models around the world have shown huge successes provided the sustainable, scalable impact is realized. With the proliferation of smart phones, clearly the future will move from the electronic platform to the mobile platform creating opportunities for mobile commerce (mCommerce), mobile business (mBusiness), and mobile government (mGovernment). In 2013, there were 200 new Internet users who ventured into the cyberspace for the first time, they represent a critical mass by themselves added to the already existing billions of users who are young passionate and technology savvy who increasingly rely on emerging ICTs to virtually do everything on a daily basis either related to their profession, study or personal life.

9 Conclusion

The digital economy promises to be a profound driver and a catalyst of change for a variety of industries. This is realized through the reach that firms can realize in terms of markets, the power and advantages created through advanced

supply chain management systems, the networks of partners and the direct access to consumers and end-users. In general, the digital economy empowered by the Internet supports different suppliers to acquire and access global customers directly through infomediaries and not intermediaries.

It is important that an adequate legal and regulatory environment is in place for eCommerce in order to manage the process, reduce the associated risks, and encourage an increasing number of people to change their trading and business purchasing habits converting from dealing in the marketplace to the marketspace. eCommerce holds many opportunities for Egypt at the business and socioeconomic development levels. With regular emerging innovations in ICTs, more opportunities are created for competition on a global scale as well as introducing new business processes, trading communities, and creating new revenue streams. Improving the economic stature of many people will undoubtedly also have concrete implications on the social stature of many. For Egypt, the digital economy represents an opportunity to keep pace with the developed world and to leverage its developmental plans; however, there are a number of developments that need to take place to transform the above-mentioned challenges into opportunities. Such developments would have to be carried out by the government in collaboration with the private sector and the civil society and would have to be on a nation-wide scale to provide a horizontal approach to development in order to avoid the creation of gaps within the community and also to be able to realize the critical mass required for a successful eCommerce community to prevail.

References

1. C. Avgerou, *Information Systems and Global Diversity* (Oxford University Press, New York, 2002)
2. D. Bell, Egypt online, *PCWorld Egypt*, Dec 1998
3. J. Coetzee, The convention on the use of electronic communications in international contracts: creating an international legal framework for electronic contracting. S. Afr. Mercantile L J. **18**, 245 (2006)
4. G. Dalakian, Can Internet giants bring down the cost of connectivity (2013), www.wamda.com. Retrieved 10 Jan 2014
5. P. Datta, A preliminary study of ecommerce adoption in developing countries. Inf. Syst. J. **21**, 3–32 (2011)
6. D.M. Davis, Residence based taxation: is it up to the e-commerce challenge? Acta Juridica **1**, 161–162 (2002)
7. G. Ebersohn, The Internet: its possibilities and risks. Juta's Bus. Law **12**, 156–157 (2004)
8. Egypt ICT Indicators (2014), www.egyptictindicators.gov.eg. Retrieved 15 Feb 2014
9. P. Esselaar, J. Miller, *Towards Electronic Commerce in Africa, a Perspective from the Country Studies*, vol. 2 (Miller Esselaar & Associates, 2001), pp. 4–9
10. N. Ewelukwa, Is Africa ready for electronic commerce? Eur. J. Law Reform **13**, 3–4 (2011)
11. C. Gentile, Forging the Future of e-Commerce, *PCWorld Egypt*, Sept 1999
12. S.J. Jacobs, H.E. Herselman, Information access for development: a case at a rural community centre in South Africa, in *Informing Science and Information Technology Conference Proceedings,* Manchester, England, June 2006, pp. 295–306

13. S. Kamel, Information Superhighways, a Potential for Socio-Economic and Cultural Development in Managing Information and Communications in a Changing Global Environment, in *Proceedings of the 6th Information Resources Management Association International Conference* ed. by M. Khosrowpour, Atlanta, Georgia, 21–24 May 1995, pp. 115–124
14. T. Kamel, Internet commercialization in Egypt: a model from Africa, in *Africa Telecom Conference Proceedings*, Johannesburg, South Africa, May 1998
15. S. Kamel, Building the African information infrastructure. J. Sci. Ind. Res. 118–144 (1999) (Special Issue on Management of Information Technology Organizations and Beyond. New Delhi: National Institute of Science Communication, March–April) (1999)
16. S. Kamel, Electronic commerce in Egypt, in *Managing Web-Enabled Technologies in Organizations: A Global Perspective*, ed. by M. Khosrowpour (Idea Group Publishing, Hershey, 2000), pp. 210–232
17. S. Kamel, Egypt's ongoing uprising and the role of social media: is there development? J. Inf. Technol. Dev. (JITD) **20**(1), 78–91 (2014)
18. S. Kamel, M. Hussein, The development of e-commerce: the emerging virtual context within Egypt. J. Logist. Inform. Manage. **14**(1–2), 119–127 (2001)
19. S. Kamel, M. Hussein, The emergence of eCommerce in a developing nation – the case of Egypt. Benchmarking – An Int. J. **2**, 146–153 (2002)
20. G.S. Kirkman, C.A. Osorio, J.D. Sachs, The networked readiness index: measuring the preparedness of nations for the networked world. The Global Information Technology Report (2002)
21. Ministry of Communications and Information Technology, *The Future of the Internet Economy, a Statistical Profile* (Ministry of Communications and Information Technology, Cairo, 2013)
22. S.M. Muniafu, E. van de Kar, J. van Rensburg, An environment for e-commerce and ICT enabled logistics services in rural areas of transition countries: the case of South Africa. in *ACM International Conference Proceedings,* vol 113 (August 2005), pp. 393–399
23. M. Odedra-Straub, Enforcement of foreign technology on Africa: its effect on society and the use of information technology. Afr. Technol. Forum **5**, 138–148 (1992)
24. M. Odedra-Straub, E-commerce and development: whose development? Electron. J. Inf. Syst. Dev. Countries **11**(2), 1–5 (2003)
25. Organization for Economic Co-operation and Development, Development Co-operation Report, Paris (2010)
26. E. Schmitt, Editorial, *PCWorld Egypt*, Nov 1999
27. R. Sorieul, The United Nations convention on contracts for the international sale of goods (CISG) as a set of uniform rules for electronic commerce. Bus. Law Int. **1**(4), 380 (2000)
28. D. Straub, M. Keil, B. Brenner, Testing the technology acceptance model across cultures: a three country study. Inf. Manage. **33**, 1–11 (1997)
29. The Boston Consulting Group, *Egypt at a Crossroads – How the Internet is Transforming Egypt's Economy* (The Boston Consulting Group, Boston, 2012)
30. United Nations Conference on Trade and Development (2000), http://unctad.org/en/docs/posdtem11.en.pdf. Retrieved 15 Dec 2013
31. United Nations Conference on Trade and Development, Forward to the UNCTAD electronic commerce and development report (2002), http://r0.unctad.org/ecommerce/ecommerce_en/edr04_en.htm. Retrieved 28 Dec 2013

32. V. Venkatesh, M.G. Morris, G.B. Davis, F.D. Davis, User acceptance of information technology; toward a unified view. MIS Quart. **27**, 425–478 (2003)
33. R. Watson, M. Myers, IT industry success in small countries: the case of Finland and New Zealand. J. Global Inf. Manage. **9**, 4–14 (2001)
34. World Bank, *World Development Report: Sustainable Development in a Dynamic World* (The World Bank, Washington, 2003)

Research on ICT Development in Africa Countries Focusing on Benefits, Trends, Challenges, and Solutions

Agnes Owuato Odongo

Abstract The ability to enhance socio-economic development through information communication technology (ICT) is a subject of an ongoing debate. Developed and some developing nations have made remarkable progress in their use of ICT to facilitate socio-economic development but Africa still lags behind in its ICT-based development agenda. The returns on Africa's efforts at using ICT to leapfrog its development are not as promising as those of the countries in the Asia-Pacific region. This chapter reviews Africa's research development on ICT focusing on benefits, trends, challenges, and solutions in an increasingly information age and knowledge-based global economy. It outlines the roles of knowledge and IT in addressing, benefits, trends, challenges, and solutions in IT development in Africa. It addresses polices that Africa must adopt to enhance the process of reducing the marginalization gap in the emerging global information world. The paper concludes that policies and strategies that seek to realize ICT opportunities for development should among other requirements; focus on generating clear ICT visions, be action-oriented, and have measurable and achievable targets.

1 Introduction

The world is undergoing an information technology revolution that has significantly changed many aspects of people's lives from education, industry, economy, and politics to entertainment. The impact of IT revolution on Africa continent may be considered in two ways. The first way may consider Africa to incur technology deficit and increase in the gap between it and the industrialized world. The second way may consider IT to help Africa reduce the income gaps between the rich and the poor. The question is "can Africa have adequate access to global information infrastructure and Information Technology (IT) age." This question must be asked because starting from an initial position of poverty, Africa cannot finance the investments in information infrastructure and computer hardware and software required to access IT age.

A.O. Odongo (✉)
Kenya electricity generating company limited, P.O Box 79439, 00200 Nairobi, Kenya
e-mail: agwuato@gmail.com; aodongo@kengen.co.ke

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_8

Therefore, Africa risked marginalization of the global economy with severe competitive disadvantages, for its goods and services. However, IT can provide means for Africa to turn its disadvantages into advantages by adjusting to the new ways of doing business, and building needed IT infrastructure. Africa is facing a complex mix of development challenges that need to be tackled in order to accelerate economic growth and reduce widespread poverty. As new Information Communication Technologies (ICTs) continue to drive rapid globalization, Africa countries have a unique opportunity to harness the power of ICTs to address development challenges effectively.

The pervasiveness and significance of ICT throughout the economy and society cannot be overlooked. The former UN Secretary-general Kofi Annan, stated that, "If harnessed properly, ICTs have the potential to improve all aspects of our social, economic and cultural life, and ICT can serve as an engine for development in the twenty-first century" [8]. What to ask is, how can ICTs policy challenges be addressed at national and international levels, and in all sectors? Adomi [2] declared that most developed countries have observed significant changes in ICTs.

However, a wide gap exists between developing countries, notably African countries, and developed ones in terms of the contribution of ICTs to the creation of wealth. If measures are not taken to make ICTs affordable and easy to use, access to them will be insignificant in developing countries [2]. The digital divide is caused by inadequate infrastructure, high cost of access, inappropriate policy regimes, network telecommunication provision inefficiency, language divides, and lack of locally created content [42].

The high cost of access to telecommunication services is an impediment to access ICTs in Africa. This is because IT has not effectively been integrated in the development agenda of most countries as reflected in the lack of ICT policies [42]. Africa has 13 % of the world population, but has very small percentage of world telephone lines and small percentage of Internet connectivity measured in terms of number of Internet hosts and Internet users [46].

In regard to the above sentiments this chapter focuses on the reviews of Africa's research on ICT development focusing on trends, challenges, and solutions in an increasingly information age and knowledge-based global economy. It outlines the roles of knowledge and IT in addressing trends, challenges, and solutions in IT development in Africa. It addresses polices that Africa must adopt to enhance the process of reducing the marginalization gap in the emerging global information world. The chapter is organized in nine sections. Section 1 is the introduction of the chapter. Section 2 is the related work. Section 3 is the literature review. Section 4 examines trends. Section 5 examines the development challenges facing Africa. Section 6 examines the benefits or the role the technology could play in overcoming the challenges. Section 7 outlines the policies that would need to be adapted by Africa to its information accessibility. Section 8 examines the initiatives to be implemented by Africa so as to overcome the problem of marginalization. Section 9 outlines the Action plan for ICT Solution Implementation Success. Section 10 is the conclusion.

2 Related Work on ICT for Development

The propagation and use of ICTs presents opportunities and challenges both to the mature service-oriented economies like Europe, Australasia, and North America and to developing African countries [19]. Current technology and market trends fall into four major groups namely: (a) Convergence that offers seamless access to a full range of multimedia resources (b) Internet that transforms business through e-commerce (c) Wireless and satellite that allows go-anywhere technologies with lower costs, and (d) Privatization and liberalization which are the vehicles that accelerates and facilitates technological advance and access to the wide range of options.

ICT and socioeconomic development (ICTD) has been carried forward by researchers. Reflect on an econometric analysis of the impact of mobile phones on national GDP suggests that greater mobile phone penetration contributes to growth in GDP [64]. The analysis is based entirely on three bodies of data: the World Bank's World Development Indicators, the International telecommunication Union's World Telecommunication Indicators, and the World Bank's Governance Indicators. ICTD research involves a consideration on human and societal relations with the technological world and specifically considers the potential for positive socioeconomic change through this engagement.

ITU estimates that the world had 1.3 billion fixed telephone lines or 19 per 100 inhabitants and that almost a quarter of the world's 6.7 billion people were using the Internet [33]. Despite the recognized potentials of ICTs for alleviating poverty, still they are not equally accessible, leaving the poorest people behind [58]. There is a set of interrelated and unfolding factors influencing the field of ICT and its role in development [16].

In 2007, mobile penetration was 28 % in Africa and 38 % in Asia, while in other parts of the world such as the American continent, Oceania and Europe, it was 72, 79, and 110 % respectively. The gap is much wider when access to the Internet is observed: less than 5 % of the population in Africa and less than 15 % of people in Asia use the Internet, while 43 and 44 % of the population use the Internet in Europe and in America [33].

ICTs has proven successful in the provision of services such as banking and health and the creation of knowledge networks between universities in India and Africa to support open, distance, and e-learning institutions. They have also proven useful as a source of multimedia entertainment and edutainment, providing information that raises awareness regarding health issues such as AIDS [58]. The use of ICTs like fixed phone lines, mobile phones, access to radio, television, and mobile banking services has improved the livelihoods of poor people living in rural areas of developing countries [56, 58].

Many developing countries and especially their poorest inhabitants still do not have access to the benefits of the information society. Attempts to remedy this have failed because developing countries have introduced information systems

"without explicit consideration of the concept of development being advanced or the processes through which it may be achieved" [40].

Research in the relationship between ICTs and development focuses on how to maximize benefits of adopting ICT to meet development purposes [13]. Harris [26] sees ICTs as an ingredient to solving poverty. However, Brown and Grant [13] warned the academic community that to accomplish this, researchers must know the dichotomy existing between researching ICT in developing countries and for development.

The technological revolution that is transforming economies and societies into information economies and information societies meets with many obstacles in developing countries and is in danger of widening the "digital divide." Conversely, the technological advances provide the opportunity for developing countries to jump to a new paradigm before problems of delivery have been solved by traditional means, both in technical and economical terms. However, realizing this potential calls for political will and support and for solutions to four key obstacles namely: (a) access to technology (b) affordability and financing (c) inappropriate regulatory frameworks, and (d) shortage of knowledge and skills to develop and implement ICT-based systems.

African leaders have a pressing duty to eradicate poverty and to place their countries, both individually and collectively, on a path of sustainable growth and development and, at the same time, to participate actively in the world economy and body politic. The poverty and backwardness of Africa stand in stark contrast to the prosperity of the developed world. The continued marginalization of Africa from the globalization process and the social exclusion of the vast majority of its peoples constitute a serious threat to global stability.

3 Africa Development Literature Review

3.1 New Economic Partnership for Africa's Development

One of the key attempts towards a collective African vision is the New Economic Partnership for Africa's Development (NEPAD). Banard and Vonk [10] report that "53 countries have been urged to implement ICTs in three crucial development arenas: education, health and trade." While NEPAD and other initiatives have contributed to the provision of ICT infrastructure with positive results as seen in the growth of Internet users, the disparities in development across Africa are enormous. Oyedemi [50] observes that policy makers and national governments in Africa are faced with the challenge of developing appropriate policies to enhance the universal diffusion of, and access, to ICT services while adopting a holistic approach taking cognizance of the social, cultural, and political needs of the community.

3.2 Economic Commission for Africa and Africa's Information Society Initiative

May 1995 marks the 21st meeting of Economic Commission for Africa (ECA) conference of Ministers, with 53 African Ministers of Social and Economic Development and Planning, participating and adopting Resolution entitled "Building Africa's Information Highway" [4]. In response to the above resolution, ECA appointed a High-level Working Group on ICT in Africa to draft an action framework to utilize the ICT to accelerate the social-economic development of Africa and its people. The outcome of the Group's work is the document entitled Africa's Information Society Initiative (AISI) which was adopted by all of Africa's Planning Ministers at the subsequent meeting in May 1996. The purpose for adopting AISI action framework was to form and develop a National Information and Communication Infrastructure (NICI) plan in every African country, driven by national development priorities [25]. The initiative proposed cooperation among African countries to share experiences of successes and to avoid failures. The countries that have begun developing in-depth national information and communications infrastructure plans include Benin, Burkina Faso, Cameroon, Kenya, Rwanda, Comoros, etc.

Overview

Okpaku [47] noticed that numerous events assisted to advance the focus on ICT Development and ICT for Development in Africa. In October 2002, the United Nations ICT Task Force held its third meeting after its inauguration in November 2001. The objectives were to: (1) provide status of current initiatives in regard to ICT development in Africa, (2) identify key issues that could have lasting impact on ICT development in Africa, (3) identify potential areas of collaboration with the NEPAD e-Africa Commission to support NEPAD's total objectives, and (4) identify areas where Task Force could assist mobilize significant support, promote, and partner with NEPAD and its e-Africa Commission, for lasting positive impact on ICT Development in Africa, within the vision and self terms of reference.

During 2002 Summit the Heads of State of the G8 industrial countries authorized the program and Implementation Plan of the NEPAD, the strategic development initiative of the continental organization, and the Organization of African Union (OAU). At the summit, the African Union (AU) succeeded the OAU; the NEPAD program was officially adopted by the new continental organization, with NEPAD as an organ of the AU.

In 2003, the World Summit on the Information Society (WSIS) adopted a declaration of principles aimed at enabling everyone to create access, utilize, and share information and knowledge. In this declaration, member states affirmed their commitment to aggressive investment in ICT for social and economic development [45].

The G8 Heads of State also adopted their own parallel program, The G8 Africa Plan of Action in 2002 G8 Summit in Kananaskis, Canada, to support Africa's initiative. The G8 Africa Plan of Action emphasized on support for ICT Development in Africa, and commits the member states to provide support to enhance Africa's ability to develop ICT capacity as well as to take advantage of the enabling capacity of information and communications technologies and applications in her drive for comprehensive development. The G8 Africa Action Plan undertakes to assist Africa to create digital opportunities by: (1) encouraging the Digital Opportunity Task Force (DOT Force) International e-Development Resources Network to focus on Africa, and supporting other DOT Force initiatives that can help to create digital opportunities, each building, wherever possible, on African initiatives already underway, (2) working towards the goal of universal access to ICT by working with African countries to improve national, regional, and international telecommunications and ICT regulations and policies in order to create ICT-friendly environments, (3) encouraging and supporting the development of public–private partnerships to fast-track the development of ICT infrastructure, and (4) supporting entrepreneurship and human resource development of Africans within the ICT Sector in regard to:

- Supporting African initiatives to use ICT to address education and health issues.
- Supporting African countries to increase access, and use of ICT in support of governance, by aiding the implementation of national e-strategies and e-governance initiatives to increase efficiency, effectiveness, transparency, and accountability.

The UN ICT Task Force participated in promoting ICT development in Africa, in collaboration with the Southern Africa Development Community (SADC), government institutions, and the African individual and institutional private sector. The World Economic Forum (WEF) conducted a comprehensive initiative on e-readiness in the Southern African Region and dedicated its most recent African Economic Summit.

The Task Force Digital Bridge to Africa Workshop

In 2002, the UN ICT Task Force convened the Digital Bridge to Africa Workshop with a view to mobilizing African ICT expertise and resources abroad in support of Africa's ICT development on the continent. Co-sponsored by UNIFEM, the United Nations Fund for International Partnerships (UNFIP), Digital Partners and Gruppo CERFE, the workshop consisted of a panel discussion. The workshop attended by some 130 participants as well as a number of African ICT experts and entrepreneurs at home and abroad, resulted in the creation of the following initiatives namely (1) The Digital Diaspora Network-Africa (DDN-A) (2) AFRISHARE, and (3) The Social Venture Fund for Africa.

The Gateway Project

The Task Force has mobilized resources and windows of opportunity through its partnerships, in support of ICT development. For example, the Gateway Project of the World Bank has provided a Global Database Gateway and a window for ICT projects to respond to the needs of the Task Force. Furthermore, it provides support through its Network of Country Gateways.

4 Africa Development Trends

4.1 Telecommunication Industry

While the telecommunications industry in the United States, Canada, and Europe invested in landlines before moving to mobile phone networks, the mobile phone has effectively leapfrogged the landline in Africa. Landlines are expensive for African countries with poor roads, vast distances, and low population densities. As per the mobile phone coverage comparison, there were approximately 8.2 million fixed telephone lines in Africa in 1998, covering 1.4 % of the population. Between 1998 and 2008, a mere 2.4 million additional landlines were installed [33].

4.2 Use of Mobile

Furthermore, sub-Saharan Africa (SSA) has some of the lowest levels of infrastructure investment in the world. Just about 29 % of roads are paved, barely a quarter of the population has access to electricity, and less than three landlines are available per 100 people [33]. Nevertheless, access to and use of mobile telephony in SSA has increased drastically over the past decade. There are ten times as many mobile phones as landlines in SSA [33], and 60 % of the population has mobile phone coverage. Mobile phone subscriptions increased by 49 % annually between 2002 and 2007, as compared with 17 % per year in Europe [29]. It was expected that by 2012, most villages in Africa will have coverage, with only a handful of countries—Guinea Bissau, Ethiopia, Mali, and Somalia—relatively unconnected [54].

Nevertheless, there have been huge disparities in the geographic rollout of mobile phone coverage, prompting concerns over an intra-African digital divide [29]. In 1999, most African countries had no mobile phone coverage, and only Egypt, Morocco, Senegal, and South Africa had coverage rates of over 40 %. By 2008, however, over 65 % of the African population had access to mobile phone coverage, with 93 % in North Africa (Algeria, Egypt, Libya, Morocco, and Tunisia) and 60 % in SSA. Although 65 % of the African population had access to mobile phone coverage, the undersea cables have been installed to run around the entire SSA coastline which was completed by 2011.

In Africa, mobile cellular subscriptions in 2010 reached 45 per 100 people, ITU-D [34] and this proportion continues to rise due to very high demand and falling costs. The growth in the capabilities of mobile and personal devices is enabling and feeding on increased availability of digital materials and applications.

4.3 ICT in Education

Innovations in technology have resulted in an increased use of ICT in education worldwide. In Africa, many governments have focused on developing national ICT policies and NICI Plans to support the socio-economic development efforts and the policies concerning ICT in education.

Currently ICT for education in Africa is occurring within national, and emerging regional, policy framework that provides the basis for partnerships and donor participation [23]. To solve ICT education problems, governments decided to formulate an education policy to address the need for more awareness, as well as ensure that sound basis for ICT education and utilization is laid [23]. Governments gave priority to policy development. Most countries have chosen to develop an ICT policy that is specific to the education sector. Thus the new phase of ICT for education in Africa is occurring within national, and emerging regional, policy frameworks that are providing the basis for partnerships and donor participation [23].

4.4 African Countries Libraries

In most African countries "libraries are not linked to the Internet and to take advantage of the potential provided by Internet facilities and use them in the provision of library and information services". This is because most African countries have: inadequate funding; inadequate telecommunication infrastructure and high cost of telecommunication facilities; and a general shortage of skilled IT human resources in libraries [18]. For African countries to become part of the GI infrastructure, libraries require sufficient funding to support their activities. Funds are required for purchase and installation of information technology, establishment of local electronic networks and connection to external networks, conversion of materials to electronic format, and training of library staff in various information technology skills. In most SSA, the high cost of access to telecommunication services is an impediment to access to ICTs [22].

Libraries in Africa continent represent one area that has experienced ICT revolution. The fact that we live in knowledge-based society with the need for universal access has made it necessary for the library to redefine its role and mode of service delivery. Emphasizing on the relevance of ICT to effective modern library management and services, Aina [3] asserted that: ICT has radically transformed most of the services provided by a library. ICT is heavily utilized in the storage, processing, and dissemination of information.

5 Africa Development Challenges

5.1 Overabundance Initiatives

Okpaku [47] noticed that the challenge to the effective pursuit of ICT development in Africa is the overabundance of initiatives, which threaten to overwhelm Africa's absorptive capacity. Numbers seem to take priority over significance in a response, which is not inconsistent with the situation in other development efforts. In 1998, the Bureau for Telecom Development of the ITU convened a meeting in Rabat, Morocco, to try to coordinate and systematize some of these initiatives. Amongst those participating in this meeting, besides the ITU, were IDRC, Bellanet, UNDP, and several African institutions, including CSIR and the Telecom Africa Corporation. While the partnership amongst the leading development initiatives, under the umbrella of the UN ICT Task Force, forged a measure of coordination, the impact of which in terms of effectiveness or achievement is yet to be realized.

5.2 Imbalance Perception of Critical Priorities

According to Okpaku [47] a major part is intractable challenge of creating an African-defined agenda, with all the benefits of knowledge, experience in situ, compelling demands of internal self-actualization, and ownership with its implicit direct accountability and responsibility to the African peoples. A challenge is the emerging possibility, albeit incipient, that there might be a genuine imbalance in the perception of critical priorities in the urgent effort to build lasting and self-enhancing ICT capacity in Africa, with its collateral direct impact on clear and measurable social, economic, cultural, intellectual, and systemic transformation of Africa and the African condition into a comfortable partner in the global dispensation.

Okpaku [47] observed that it is impossible to undertake strategic African capacity-building in the ICT sector without first determining the extent, scope, and quality of the expertise Africa already has and which it can deploy as its first line of attack in trying to achieve the quantum development. This is necessary to move in leaps and bounds to catch up with and join the ranks of global competitive ICT development and capacity.

5.3 Africa Structural Impediments

Africa's inability to harness the process of globalization is a result of structural impediments to growth and development in the form of resource outflows and unfavorable terms of trade. At the same time, we recognize that failures of political and economic leadership in many African countries impede the effective

mobilization and utilization of scarce resources into productive areas of activity in order to attract and facilitate domestic and foreign investment.

The continent has not fully responded to call for giving ICT necessary priority in the national development despite a large number of international funding agencies with ICT initiatives in Africa. However, the situation is changing rapidly in recent times, especially after the African Development Forum (ADF'99), held in Addis Ababa in 1999, with the theme, "Challenge to Africa of Globalization and the Information Age."

HIV/AIDS illnesses, lack of access to Internet and telecommunication facilities, and lack of good ICT policies which encourages ICT inflow, have been powerful obstacles for the social and economic growth of rural Africa. Many private sector firms and civil society organizations with an interest in supporting education programs and technology initiatives in Africa have expressed similar frustrations [23].

5.4 Governance Challenge

The governance challenge of Africa is particularly acute. In a number of states, the capacity of the public sector to deliver basic services is exceptionally weak. Citizens do not trust public institutions and service providers, as many of these institutions remain opaque and fail to deliver on basic services. Weak governance, low public sector capacity, and poor service delivery operate in a vicious cycle. State performance is disappointingly weak, putting progress on development goals at risk. Results of interventions on the supply side alone through e-government are disappointing in many countries, and so scaling up demand for good governance (DFGG) is logical and complementary. Capacity to use data through the application of social accountability and mobile communication tools is increasing.

Political participation, citizen voice, and social accountability remain weak. Service delivery and underlying institutions and processes are often weak, or corrupt and nontransparent. Fighting corruption remains a major challenge and source of citizen distrust and dissatisfaction. Most Middle East and North African countries grapple with bottom rankings in the Human Development Index (HDI) and Corruption Perception Index (CPI). Corruption levels are perceived as highest at key service ministries: health and education. Government information remains closed to most citizens, businesses, and other stakeholders. With the exception of Tunisia, none of the countries have initiated a "freedom of information act" or created the enabling processes and practices in support of open data and open government initiatives [43].

5.5 Challenges for African Countries to Create Jobs

There is a challenge for African countries to create jobs through inclusive and sustained private-led growth. The networks of privileges and barriers to new businesses discourage many would-be entrepreneurs [67]. Excessive licensing and unaccountable public agencies deter investment and are major source of corruption. Privileges, unequal playing fields, noncompetitive public procurement, and lack of information about the rules of the game undermine trust, hope, the credibility of reforms, and inclusive and sustained growth. Business-entry procedures are uncertain and time consuming. Small and medium enterprises face imperfect information about new technologies and export markets. Business development services are underdeveloped for SMEs. The vast potential of the private sector, particularly SMEs, remains untapped.

Manufacturing and investment is declining, and the GDP is lower than that of all other developing regions, except SSA [67]. The region is losing the race in technological upgrading which is the key for competitiveness, structural transformation, and sustained growth. With very few exceptions, the ICT revolution has yet to be leveraged to support export, diversification, knowledge diffusion, innovative entrepreneurs, and SMEs. Yet, these challenges are not insurmountable, as the region has substantial endowments and opportunities.

5.6 Gaps in ICT Development

There are gaps in development ICT in Africa. A greater presence of advancement in terms of the use of ICTs has been noted by Banard and Vonk [10] in the peripherals of the Southern and North African States, followed by the Eastern and West African States. Most of the Central African states have been lagging behind.

According to Vukanikids-DTI [59], the developmental gaps are compounded by the problem of the digital divide, due to lack of access to ICTs, challenges of inadequate pools of skilled persons, and the use, maintenance, and rapid obsolescence of the ICTs because of continuous technological innovations and development. To remain relevant, universities worldwide and Africa in particular are under pressure to redefine mission and review curricula to produce African graduates with “global” skills [68].

People in rural areas of Africa are afraid to use computers. Some of the challenges faced by the ICT industry in rural Africa include: (1) rural resistance to change, (2) rapid technological changes, (3) limited bargaining power to access technology, (4) high cost of Internet access and equipment, (5) lack of understanding of ICT sector, (6) lack of access to low cost devices for telecommunications, (7) scattered population in rural areas of Africa, (8) availability of resources like computers, (9) illiteracy, (10) lack of interest in technology adaptation, (11) lack

of infrastructure, (12) lack of participation in international meetings, seminars and themes related to ICT for development, (13) rapid changes in software and communication industry, (14) lack of a good ICT policy.

Cuban [20] noted that the shortage of teachers will continue to be the main challenge for teacher policies in the near future. This will be the case worldwide, although reasons vary demography, labor market trends, the impact of HIV/AIDS, and so on.

5.7 Lack of National ICT Policy

Lishan et al. [1] says that a country's National ICT Policy should outline the vision of the country is in relation to ICT, and the vision for each major social or economic sector. Challenges to be taken into account when considering the use of ICT in education include: Trucano Michael [55], Neil [44] (1) the absence of comprehensive policies that enable and support interventions, and which are themselves supported by clearly defined and resourced strategies for implementation at national level as well as at the level of educational institutions, (2) the lack of financing and prioritization of ICT investments as a barrier to effective ICT use, (3) limitations in the infrastructure required to support the use of ICT in education. The two major factors are limited access to power (about 60 % of the population in Africa lacks domestic access to electricity) [37] and the lack of affordable and reliable Internet access (less than 12 % of the population in Africa are believed to be Internet users by March 2011) [35], (4) a lack of capacity at all levels to integrate and support the use of ICT effectively in education, (5) a lack of necessary ICT skills amongst teachers and of specific training to enable them to use ICT effectively in the classroom, (6) limited supply of appropriate content, including learning materials and learning support tools, (7) lack of accurate, comprehensive, up-to-date data on education, and (8) the tendency of ICT to accentuate social, cultural, and economic disparities.

There are general issues of the contribution of ICTs to development. For example, Madon [39] examined the use of the Internet in sectors such as health and education, and in domains such as economic productivity and sustainable development. Silva and Figueroa [53] emphasized the importance of a standards and telecommunications infrastructure in supporting ICT applications. Sayed and Westrup [52] looked at the role of enterprise resource planning (ERP) systems in bringing together networks in a globalized world. Aman and Nicholson [6] and Korpela et al. [36] emphasized the specific difficulty in non face-to-face communication when working across cultures. This is about sensitizing collaborators in cross-cultural offshore development projects to issues and problems which must be handled [63].

Bringing a technology to a new local context involves some implicit elements of cultural transfer and mutual learning. Macome [38] and Korpela et al. [36] concluded that the local context is crucial in the implementation process, and that it is essential to involve local stakeholders in the entire process. D'Mello [21] and

Korpela et al. [36] focused on local adaptation related to new ICTs by addressing the adaptation of people in contexts such as global software outsourcing.

Mosse and Sahay [41] and Korpela et al. [36] argued that there is need to create counter-networks, using a term drawn from Castells [14, 15] to the existing dominant networks of human and nonhuman actors operating within deep-rooted sociocultural structures Metcalfe and Joham [46]. Korpela et al. [36] argued that modern technology, such as UHF citizen band radio, can be highly effective in supporting knowledge exchange between groups with strong oral traditions. Okunoye and Karsten [48] argued that there is need for specific use of technologies such as email, databases, and telecommunications for knowledge creation, application, and storage, to generate some implications for future practice. Waema [60, 61] noted that there is need to develop ICT policies stems from the challenges facing African countries because of globalization and liberalization.

Many countries lack adequate human capacity in ICT. In Kenya, for example, most of the high-end ICT training takes place in public institutions, but these institutions lose staffs to the private sector who offer better salaries. For example at the University of Nairobi, ICT literacy courses for all students have not been started due to lack of adequate staff to deal with the large number of students (about 20,000) and inadequate computing infrastructure, especially computer laboratories, amongst other reasons.

6 Technology Can Overcome Africa Development Challenges

ICT can aid the development of African economy. Many of the sectors of the economy have not been able to grow, as they should, because of the inappropriate operating environment during policies implementation.

6.1 Introducing Economic Measures

Software Investment Opportunities

ICT applications need well-developed software for implementation. Programming language design, systems software development, and applications programs are required in the sector. Offshore software development, in particular programming, will provide opportunities for the African youth to earn foreign currency, without leaving the continent soil to look for their sustainability.

IT Parks

The African governments should invest in the establishment of ICT Parks. Infrastructure facilities should be provided in such a park and the private sector should be encouraged to establish ICT factories, training centers, and other ICT undertakings in the park. Economic incentives shall be provided in order to encourage the private sector investment in the park.

Mobile Internet Units

The African governments should invest in Mobile Internet Units (MIUs). The units are devised to aid the teaching of ICT in remote places all over the continent. The MIU is made up of a bus, whose interior is restructured into an Internet center, with the provision of computers, local area network, printers, and photocopy machines.

Mobilizing the Agricultural Value Chain

Agriculture sector uses specialized mobile services that provide localized information about price, weather and climate, pest control, cultivation practices, and agricultural extension services. The mobile services are to be used in agriculture and remote and satellite technologies that assist in food traceability, sensory detection, and status updates from the field.

m-Health

Mobiles can transform and enhance the delivery of primary and secondary healthcare services. It reviews on the ground implementations of medical healthcare applications to draw key conclusions on how m-Health can best be implemented to serve the needs of people, as well as identify the major existing barriers to be overcome.

Mobile Money for Financial Inclusion

The use of mobile as a general banking platform and critical infrastructure underpinning other economic sectors. It shows the benefits and potential impact of mobile money, especially for promoting financial inclusion. The growth of mobile money services, the barriers, and obstacles hindering their deployment, and emerging issues that the industry will face over the coming years needs to be looked into by all African countries.

Mobile Entrepreneurship and Employment

The potential of Mobile banking as a platform for employment for employment, not solely in terms of job creation, but also the ability to facilitate entrepreneurship. This is more so in populations otherwise disconnected from the economy, encourage development of transferable technical and business skills, match jobs with workers, and create opportunities for micro work.

Making Government Mobile

Governments have begun to embrace the potential for mobile technology to put public services literally into the pocket of each citizen, create interactive services, and promote accountable and transparent governance.

Policies for Mobile Broadband

Policy recommendations for expanding Mobile banking platform in terms of broadband infrastructure are expected to address the key bottlenecks of both supply and demand sides of mobile broadband.

Standardization Versus Localization

The current era has seen the development of global systems and approaches which aim to transfer best practices and procedures between different contexts and countries. But, there is a strain in developing these systems between wishing to standardize for efficiency and comparability purposes, and the difficulty of imposing the same standards on different local contexts. Braa and Hedberg [12] described the development of software and the setting of standards in an action research project concerned with improving health which started in South Africa, but the approach has been used in a several developing countries.

Alignment of Actors in Networks

A popular theory in the Information System (IS) literature currently, in developed and developing countries, has been actor-network theory (ANT). ANT offers a way of conceptualizing technology as one of the "actors" in the actor-networks. Braa and Hedberg [12] referenced the theory as a way of conceptualizing actor-networks involving heterogeneous elements such as people, organizations, software, and standards. Rolland and Monteiro [51] described the implementation of a complex information infrastructure in a global maritime shipping organization with offices in more than 100 countries, including many developing countries.

Particular Technologies

Particular technologies like ubiquitous Internet and ERP systems are the current fashionable technological debate revolving around the use of open source and free software, and indeed it is clear that some developing countries are taking strategic initiatives to encourage the open source route. African countries should embrace the use of open source and free software. Braa and Hedberg [12] states that their South African origin health information systems program (HISP) software is open source.

6.2 Looking Ahead

We need an abstract view of the role of ICTs in Africa countries contexts and how to study it to classify existing work, identify gaps, and suggest future opportunities. ITD [32] focuses on existing studies which are strong on addressing the specific challenges, gaps, and opportunities for the future.

The Development That ICTs Aim to Contribute

ICTs can promote development. There is potential for its further use in Africa contexts, for example in research on public sector organizations where institutional arrangements are often crucial. There is little work done in IS in African countries in economics development. Many micro-level projects and approaches involving ICTs have important economic elements that are often ignored in the current literature. These issues have been made pertinent through the debates on Business Process Outsourcing and economic implications it has upon unemployment levels and foreign exchange reserves [28].

The Key Studied Issues Related to ICTs

There are some important but neglected topics. There is literature on major issues and topics concerning IS in developing countries like Africa countries. These issues comprise of local adaptation and cultivation of ICTs, standardization versus localization, and detailed studies of particular technologies. In developing countries, there are issues that are considered important, but which have received less attention in the literature. For example, research has been conducted on the use of ICTs in developing countries focusing on health, public administration, and education, a major problem remains the issue of scalability. Literature contains little discussion on how to tackle scalability. Another neglected topic is sustainability that focuses on how ICT-based projects can be sustained with appropriate resources.

Exhaustive Studies of Other Technologies

Exhaustive studies of particular technologies, and the detailed way in which the hardware, software, and system configurations interact with the social, economic, and cultural context. A number of other areas have received limited attention. An example is the area of e-government, where the precise way in which the particular software, hardware, and related networks are set up for a specific application is crucial in determining application failure or success.

Open Source and Free Software

The distinctions between the two, and subdivisions of these, are important issues in influencing whether and how a particular African country can use a specific piece of software. Licensing agreements are variable and free software, for example, may give free use but no access to the source code. Open source software may require that any future product based on the software be made available to others, or in some cases not.

Large-Scale Technological Infrastructure Investments

Despite a general improvement in telecommunications coverage in the developing countries over the last decade, connectivity and quality in poorer areas and countries often remain inadequate. Why is this? And what can we do about it? The Internet provides a second technological example where more detailed research is needed. Writers in the richer countries sometimes speak of the networked society where everyone is connected online. This is certainly not the case in the developing countries, and even within poorer groups in the rich countries of the world. The digital divide is a question of access and of educating people to be able to use the resources of the Internet effectively.

Provision of ICTs in Conjunction with Roads and Electricity

It is not uncommon in the poorer countries of Africa to find expensive computing equipment lying idle due to an inadequate power supply, or the unavailability of spare parts or technicians due to impassable roads [41]. Korpela et al. [36] described how the large physical distances and lack of adequate transportation impede the flow of health information in Mozambique.

Society-Based Critical Issues

*S*ociety-based critical issues like HIA/AIDS demand attention. By trying to target antiretroviral drug treatments, good patient data and records are needed, something which is often lacking in the poorer countries where the disease has its worst effects. Research is needed to address how best to tackle the informational aspects of HIV/AIDS, and link these systems with policy processes regarding drug rollout and education campaigns [17].

The Theoretical and Methodological Stance

Methodologically, the existing literature is stronger than a decade earlier, with comprehensive studies and those dealing with interconnected levels of analysis being more common. With respect to theory in future studies, suggestion is that there is need for critical studies in the academic sense of the term Orlikowski and Baroudi [49]. Africa as a continent having countries in the level development is normally deeply intertwined with issues of power, politics, donor dependencies, institutional arrangements, and inequities of all kinds.

The Level and Focus of Analysis Adopted

Diverse levels of analysis can be identified when researching ICTs use in Africa that is made of developing countries. These may include individual, group, organization, sector or national, and international [62]. The neglected level of analysis is the individual level. Yet, issues of shifting identity for example, as noted for Indian software engineers in D'Mello [21], are vital to improved understanding of a particular phenomenon in the field.

7 Policies for Information Accessibility

7.1 Policy

In the strategies for finance and funding, it is stated that government shall provide incentives to investors to enable them grow rapidly and efficiently. Policy statement should be terse, straightforward in order to achieve the set goal.

Lack of Proper Planning of Policy

There is regrettably lack of proper planning of policy with constructive analysis and consultation of people in the field for the realization. There is a rush to borrow what exist in the developed world without considering the different circumstances and peculiarities of operation. Successful policy implementation in African countries requires a great deal of political will. For example, in a bid to position Thailand in a technology driven and interconnected world, Thailand government created a separate ministry for Information and Communication Technologies [9]. In most African countries the problem of insufficient and unreliable data is a perennial one. The quality and reliability of the statistical data on which policies are based are weak and non-existent.

Alabi [5] made a case for the need for an information policy in Africa embracing information technology and other related policies. World Bank noted that developing countries are facing challenges in adapting policies and regulations to rapid changes in technology and market structure Farrell et al. [24]. The most critical laws and policies arise in the following areas: (1) policies that directly impact ICT innovation and investment; (2) policies that develop people's capacity to utilize ICTs effectively; (3) policies affecting entrepreneurship and Research and Development (R&D); and (4) policies affecting telecommunications infrastructure and access.

Constrained Policy-Making Institutions

Policy-making institutions in Africa remain constrained by lack of access to timely information, and limited capabilities for policy analysis, data mining, information sharing, and knowledge management. They suffer from cumbersome data collection and fragmented reporting systems combined with overload of unprocessed data and "information poverty."

7.2 Regulatory Matter

According to Okpaku [47] national regulatory authorities are forming sub-regional groupings, such as the West African Telecommunications Regulatory Authority (WATRA) and the Telecommunications Regulatory Authority of Southern Africa (TRASA). Taking advantage of common attendance at various forums, such as the African Regional Preparatory Conference for the World Summit on Information Society in Bamako, Mali, in June 2002, African experts, officials, and sector entrepreneurs are holding ad hoc meetings to pool their resources to pursue common interests. In this regard, moves are underway for the creation of an association of African Telecommunications Regulatory Authorities, with the support of the Bureau for Telecommunications Development (BDT) of the ITU.

Infrastructure

There are several regional infrastructure initiatives being taken by Africans in the ICT sector. The initiatives include:

The Sat-3/WASC/SAFE Undersea Optical Fiber Cable Network

To show of continental solidarity, cooperation, and strategic common purpose, several African countries, in the euphoria of having achieved the most singular strategic objective of the OAU, undertook to join forces to build a major undersea optical fiber cable to directly link many African countries, and link them to Europe and Asia. Led by Telkom S.A., the 36 participants built the Sat-3/WASC/SAFE cable network. The 28,000 km cable, which cost US$650 million, was launched in Dakar on May 27, 2002.

The RASCOM Satellite Project

RASCOM, the Regional African Satellite Communications Organization to which most African governments belong, have undertaken to build the RASCOM Satellite, in collaboration with Alcatel. Designed to have a footprint covering the continent, it is intended to support affordable access to ICT resources for Africans anywhere on the continent, especially rural populations.

The Comtel Project

Comtel is a regional project, undertaken by the member states of COMESA, to build an optical fiber grid to interlink their national networks.

7.3 Policy and Strategy

The African Connection and the Ministerial Oversight Committee

The African Connection was created by the African Ministers of Communications during the ITU African Regional Conference in Johannesburg in 1998, to serve as an institutional framework for the coordination of telecommunications development ideas and capacity-building, for those with regional scope. The African Connection is supervised by the Ministerial Oversight Committee of African Ministers of Communications.

Table 1 Development of national ICT policies 2000–2007

Status of national ICT policy development by country	2000[a]	2005[a]	2007[b]
Policy in place	13	28	36
Policy under development	10	15	12
No development underway	30	10	5
Total	53	53	53

[a]Source: UN Economics Commission to Africa (UNECA), http://www.uneca/aisi/nici
[b]Source: ICT in education in Africa survey report

The African Telecommunications Union

The African Telecommunications Union (ATU), a reconstitution of the Pan-African Telecommunications Union (PATU) by the African Ministers of Communications, is the de facto African regional telecommunications counterpart of the ITU. ATU, which also reports to the Ministerial Oversight Committee, serves as the organ for the systematic pursuit of telecommunications development in Africa.

The African Advisory Group on ICT

In the area of mobilizing Africa's global expertise at the cutting edge for Africa's ICT Development, the African Advisory Group on ICT (AAG-ICT) has played a critical role. The AAG-ICT consisting of 12 eminent African ICT experts from around the world met behind closed doors twice yearly to provide confidential high-level advice to African Ministers of Information and Communications on strategic, policy, and regulatory issues, with no holds barred. The AAG also works in close liaison with the African Connection and the African Telecommunications Union, the heads of which two institutions also sit on the AAG. The new technologies can be used to harness rural development goals. According to the World Bank (1998:156) recent development thinking has been based on the assumption that markets work well enough to ensure development and alleviate poverty.

7.4 National ICT Policies

Glen et al. [23] argues that future socio-economic development should embrace the use of ICT that appears to be widely recognized by governments throughout Africa and is evidenced by the number of countries that have a national policy for ICT in place or under development. However, the policies vary. Those that have been in place for some time are more likely to focus on telecommunication technologies and their regulation and less on the importance of information technologies for development (ICT4D). More typical, however, are the more comprehensive policies such as those developed in Botswana, Namibia, Zambia, and Libya (Table 1).

7.5 Observations Made on ICT Policies

ICT Policies Act as Catalyst for ICT Policy Development in Education

Of the 48 countries that either have a national ICT policy in place or are in the process of developing one, 39 of them have education sector ICT policies and are in the process of developing them. Clearly the development of national ICT policies acts as a catalyst for ICT policy development in the education sector. In fact in some cases (e.g., Botswana, Rwanda) the development of goals and implementation plans for the education sector has been an integral part of the larger process of developing the national plan.

All ICT/Education Policies Stress Enhancing Access, Some Go Further

All policies surveyed emphasize the importance of enhancing access to ICT tools and Internet connectivity, developing ICT skills among young people and the general population, and the importance of teacher training. However, some policies stress on the need for the development of digital content, education portals, and the need for content.

Donors Play an Important Role in the Policy Process

The international donor community (e.g., UNECA, USAID, UNDP, AfDB, IDRC) continues to play an important role in the policy development process, and the impact is seen in many plans in terms of the detail where purposes, outcomes, performance indicators, monitoring and evaluation strategies, and implementing strategies are set out. The plans developed by South Africa, Kenya, Mauritius, and Rwanda are exemplary in that respect.

Infrastructure in Schools

African Ministries of Education have begun to be more proactive in coordinating and leading the development of ICT infrastructure in school systems as their ICT policies and implementation plans have taken shape. However, civil society, principally non-governmental organizations (NGOs) working with donor agencies, continues to play a major role in providing computers to schools and lobbying governments to take a leading role. However, their efforts have been frustrated by the lack of connectivity, inconsistent electrical supply, and lack of technical support services.

A Need to Change Existing ICT Policies

Meera et al. (2006) observed ICT policies for all African nations, which need to be reviewed. The ICT policies should aim at bridging the digital divide by training people in education, health, business, tourism, and government sectors who will in turn help in the social and economic development of the developing nations.

8 Initiatives to Overcome Marginalization

There is marginalized demand for connectivity in African countries. The information kiosks set in marginalized areas will assist children to Google the Internet, run training courses using computers and many more to perfect access, and ICT use.

8.1 African Initiatives

The Global Human Resource Survey of African Male and Female Experts in ICT

Are designed to create a critical database for strategic decision-making policy and access to African cutting-edge expertise, wherever it may be worldwide [47]. This anticipated database will enable African governments and their decision makers, development agencies, international organizations, and the global private sector, to draw on Africa's already vast human resource expertise to shape, formulate, man, and drive critical ICT initiatives throughout the continent.

The Telecom Africa Virtual Research Laboratory Project

To link African scientific and technological research experts around the world and their global counterparts with interest in African ICT development in a secure global Intranet to undertake research activities of specific focus on African needs and priorities [47].

Informal Sector and Civil Society Initiatives

The informal sector, consisting primarily of NGOs, plays a significant role in advancing efforts at building ICT development in Africa, especially with a view to the use of ICT technologies and applications for dealing with the scope of issues which have become traditionally associated with NGO efforts: namely, the eradication of poverty; the social, educational, and political empowerment of

the disadvantaged or erstwhile deprived, especially women, children; rural and variously handicapped persons; education; preventive healthcare and the management of illness, especially HIV/AIDS and other communicable diseases; and universal access to basic information and telecommunications services.

SMEs

Given the size of the informal component of the African economic landscape, especially with over 70 % of the population living in small and rural communities, and the unique adaptability of ICT applications to small and micro enterprises, it is not surprising that these enterprises have become increasingly active in the African ICT environment. Taking advantage especially of the online facilities of the Internet, as well as the development of telecenters, individual and small groups of African entrepreneurs are setting up a slew of businesses, from online marketing of farm products, arts, crafts, and clothing to Internet cafes and telecenters.

Industry-Based Initiatives

The ICT industry in Africa itself has gradually come around to the realization that its long-term profitability in the African market is intimately tied into the development of ICT capacity, not only in order to increase market demand, but also with respect to promoting economic and social development. This is the only way to increase the buying power of the African population, which, in turn, will increase that portion of their income which they can then invest in availing themselves of more and broader ICT products and services.

Indigenous African Initiatives

Of great importance in the ongoing drive for ICT Development in Africa is the active role of Africans and African institutions, both public and private sector, in undertaking often quite bold and innovative initiatives. The potential impact of such efforts is itself greatly enhanced by the intimate level of collaboration between sectors, government and private, and between Africans on the continent and in the Diaspora. These initiatives cover a broad spectrum of areas, from policy and regulation, to industrialization, infrastructure, software, content, development communications, and capacity-building.

The Sustainable Development Networking Program (SDNP) has been a pioneer in ensuring ICT is used for development in least developed countries as well as developing countries, through various ICT programs and initiatives:

Women's Voices

A video initiative for women empowerment and social participation was implemented in Kenya.

African Virtual University

Provides global knowledge transfer, e-learning concept for least developed countries.

NEPAD Programs

ICT initiative to launch e-schools, free Internet access, computers, e-health initiatives that focuses on least developed countries.

iConnect (AfriConnect)

Provides high speed wireless broad band Internet services and rural connectivity in Zambia, Tanzania, and Kenya.

Task Force to Support African ICT Development and NEPAD's ICT Program

The New Partnership for Africa's Development (NEPAD) is the strategic platform of the newly created African Union, the continental African political and economic institution that succeeded the OAU and which mirrors other regional institutions, such as the European Union [47]. An instrument of the African Union, NEPAD, in its short life, has established itself as a reference point in virtually all initiatives having significant components involving, Africa. NEPAD enjoys high priority with the United Nations.

Amongst the African items and events on the Agenda of the 57th Session of the United Nations General Assembly, such as the debate on the Causes of Conflict and the Promotion of Durable Peace and Sustainable Development in Africa (Agenda Item 33), the final Review and appraisal of the United Nations New Agenda for the Development of Africa in the 1990s (UN-NADAF) (Agenda Item 41), the General Assembly held a High-Level Plenary Session on NEPAD. From an institutional structure point of view, NEPAD has four main sectors, namely: (1) Infrastructure, which consists of ICT, Water and Sanitation, Transport and Energy, (2) Agriculture and Market Access, (3) Human Development, covering Health and Communicable Diseases, Education and Poverty Eradication, and (4) Capital Flows, which consist of Domestic Resource Mobilization, Private Capital Flows, ODA Reform, and Debt Reduction.

The e-Africa Commission

Positioned within the Infrastructure Sector, ICT is a major focus in the NEPAD agenda. To oversee this process, African leaders formed the e-Africa Commission to serve effectively as the ICT task force of NEPAD in pursuit of the NEPAD objectives. The focus is on "Development and International Co-operation in the twenty-first century mainly on The Role of Information Technology in the Context of a Knowledge-based Economy."

NEPAD has called for the cooperation, expertise, synergy, and the strongest commitment of support African and international institutions. This is meant to ensure successful management of the process in a way as to fulfill the enormous expectations the African people, and indeed the world, have placed on the young institution.

The Program of Action for Least Developed Countries by the Third United Nations Conference held in Belgium on the Least Developed Countries [7] is meant to halve the cases of extreme poverty by 2015 as well as promote sustainable development for the marginalized group.

The Integrated Sustainable Rural Development Strategy

ISRDS [31] held in 2,000 mirrors the views about the marginalized group development. Marginalized group development is about multi-dimensional, encompassing improved provision of services, enhanced opportunities for income generation and local economic development, improved physical infrastructure, social cohesion and physical security within marginalized communities, active representation in local political processes, and effective provision for the vulnerable.

Bhavnani et al. [11] observed that despite the increasing rural demand for relevant and timely information and market knowledge, the benefits from ICT investments have been unevenly distributed between and within countries resulting in the digital divide and information poverty. The causes of the situation, to be tackled include:

Institutional Environment Constraints

The enabling policy and regulatory environment are not conducive to stimulation of competition and increased private sector involvement in the provision of ICT infrastructure and services to rural communities. Typically, there is also a lack of well-developed and functioning institutional mechanisms to implement the policies and regulation; lack of locally relevant easy to understand content in local language; and lack of well-trained human resources to develop applications and service the end-users.

Rural Infrastructure Constraints

Rural ICT infrastructure is underdeveloped, due to the high costs of last mile connectivity in rural areas, intermittent and unreliable power, and low priority for ICT investment, due to other more pressing needs in the rural sector.

Rural Population Constraints

Rural population characteristics are not conducive to ICT absorption. Barriers include: low population density and remoteness, low levels of functional literacy, little or no basic or computer literacy, low awareness, low disposable income, poor health and living conditions, and struggle for survival.

Rural Poverty Reduction Strategy Constraints

ICT is not an area that has been well integrated in rural poverty reduction strategies: often narrowly defined as modern technologies (e.g., computers and the Internet) and the more traditional technologies (e.g., fixed line telephone, radio, and television) have not yet been fully exploited.

8.2 Barriers to Local Adoption of ICTs Through Rural Development Programs

Target Driven Projects

ICTs introduced for a specific purpose can exclude other sections of the community and institutions that have a greater role in local information and knowledge systems. For example, a program on breast feeding may "reach" the target audience of 50 % of mothers within a certain age group or area, but the extent to which the programs acted upon may be limited by the institutional context in which the ICTs and the information are controlled.

Presumption of Modernity

Much of the technology is so new to development agencies, let alone rural communities, that it is difficult to avoid equating access to modern equipment with existing social hierarchies. This will always lead to elements of "as a means of raising or reinforcing an individual's status". Those with a higher status due to education will be given preferential access on the assumption that they are more likely to treat the equipment properly and learn how to use it more easily.

Cost and Quality

Very few ICTs in use in rural development programs in developing countries have been designed for that purpose. The result is that the technology is often too expensive and too fragile to be really used freely for public access. Increasingly portable, lightweight, robust, and cheap ICTs such as mobile phones are reducing this but problems of infrastructure such as electricity and telecommunications networks need to be addressed. The development of more flexible ICTs for rural development based on low unit and running cost, high portability, and information outreach potential would greatly reduce the tendency towards restrictive control.

Encouraging Participation and Demand-Driven ICT Use

To integrate ICTs into cross-sectoral rural development strategies it is important to understand the local context into which information is to be made availed and used. Although a radio play may or may not achieve the best impact possible, it is important that local institutions and existing systems develop ways of improving information exchange using ICTs in a participatory manner. The process of learning to stimulate the demand for information from local communities is provided by first hand training and capacity-building.

Improving Social Inclusion of Isolated Populations

People living in rural and remote areas tend to be poor and socially isolated [27]. They lack information relevant to their particular situation and so have difficulty interacting with other community members. This isolation serves to reinforce their marginalization. ICT, such as radio, telephone, and email, can be of great value in bringing people together, bridging geographic distances and providing relevant information about and to the poor. The four characteristics of ICT, namely (1) interactivity, (2) permanent availability, (3) global reach, and (4) reduced costs for many, as summarized in a recent report by the Swiss Agency for Development and Cooperation (Gerster and Zimmermann 2003), have made social inclusion of the poor more feasible.

Facilitating Political Empowerment

The poor lack means to voice their needs to learn about available public services, and pressure policymakers to be responsive to their interests and demands. Increasing ICT use for developing propoor policies is a critical component to reduce poverty and sustain development. In India, Sakshi, the women's rights NGO, with help from international women's networks, received advice and technical assistance on legal issues surrounding sexual harassment. As a result, the group succeeded

in convincing the Supreme Court to establish sexual harassment guidelines in workplaces and brought the issue within the purview of human rights violations (APC 2000).

Influencing Public Opinion on Gender Equality

ICT can be an invaluable tool in changing people's attitude towards women by disseminating educational programs on gender equity. Radio Rabia Balkhi in Mazar-E Sharif, a remote northern city in Afghanistan where more than 80 % of women are illiterate, airs recorded essays and features on gender inequality issues and successful efforts to overcome these challenges. After listening to the program, some conservative family heads in the region, which had refused to let their daughters attend school, changed their minds (Constable 2003).

In Non-agriculture Jobs

Some women work in the expanding ICT industries, but data from developed countries suggests that men dominate these new opportunities, mainly because from the outset women lag men in access to ICT and skills to apply it. A study by Hafkin and Taggart (2001) indicates that women are 22 % of all Internet users in Asia, 38 % of those in Latin America and only 6 % of Middle East users. Women need access to educational resources in order to lower overall illiteracy levels and acquire skills required for the ICT sector.

ICT for Women's Education and for Women as Educators

In most societies, women are predominantly responsible for child-care, food preparation, and other household tasks. Therefore, providing women with useful information may have multiple benefits—in addition to the benefits to the woman herself, the technical and substantive knowledge gained by women is often shared with their families. Educated women do a better job caring for their children, increasing their children's chances of surviving to become healthier and better educated adults [65].

Enhancing Women's Ability to Know Their Rights to Participate in Decision Making

Women remain vastly under-represented in national and local assembles accounting for less than 10 % of the seats in national parliaments, on average worldwide [66]. ICT can empower women by enabling them to participate in public discussions

and democratic processes using multi-way communications, through electronic and non-electronic ICT networks. In Kenya, for instance, a group of women used videos to inform the decision makers about their needs and concerns.

9 Action Plan for ICT Solution Implementation Success

The way to achieve maximum benefit for ICT implementation is to have all the factors for success with no occurrence of the factors for failure. Some ICT best practices "harvested" from a review of successful applications are suggested by Clockwork. This can be viewed from an angle of "technology leapfrog" which can be achieved by appropriate technology transfer (Ifinedo 2005). The best practices in ICT are:

- Do not underestimate the complex environment in which ICT programs evolve.
- Select a project that demonstrates the greatest benefit for your target group.
- Re-skill government staff to anticipate the changes in ICT structure and roles
- Identify the right technologies.
- Determine how an organizational process fits your technology.
- Strong program and project management to implement successful ICT solutions.
- Do not underestimate the total cost of ownership (TCO) of an ICT project.

9.1 Solutions

Waema [60, 61] argues that countries should have policies that encourage all workers to be computer literate. Structure for ICT policy development and implementation is important for governments to set up national entities that champion the development and implementation of ICT policies and strategies.

ICT Human Capacity Development

In resource scarce countries, like in Africa, human resource is the main factor for economic prosperity and it could be a real asset in the digital age, so each nation should accord investing in human capacity development, in ICT, high priority. Efforts to expand capacity for ICT education and training also attract young students into the profession build indigenous ICT capacity.

ICT Literacy for All

While developing the human capital we should take into account that ICT is centered on people with appropriate skills and knowledge. The basic building block of ICT industry is the skilled and semiskilled manpower.

Cost of ICT

The cost of ICT products and services can be reduced through: Instruments to encourage private sector participation. These include bank loans with low interest rates for the education sector and zero rating duties and taxes for ICT products and services in education. This will enable the private sector to be actively involved in ICT human capacity development and also offer other ICT services.

Quality Programs and Content

The ICT programs and content should be of quality and characterized by the following:

Appropriate mix of staff complement: This is largely due to the growing multi-disciplinarily nature of ICT. In computer science, for example, a staff complement of "traditional" computer scientists is a recipe for disaster in training ICT manpower.

Standards for Quality Training

The stakeholders need to look for methods to regulate the quality of training institutions.

10 Conclusion and Recommendations

Africa has been left behind in the global trend. The world has moved from the dependence on natural resources to a new knowledge-based foundation, where the wealth and powers of nations are directly dependent on the strength of their ICT and the associated human capital. In this area, the development and successful implementation of ICT policy in African countries holds great hope for the sustainable economic development of the continent. It is the responsibility of all African governments and the private sector to transform the digital divide to digital opportunity.

The WSIS raised awareness of African policy makers and stakeholders on ICT issues at an unprecedented level, hence increasing expectations on what they can get out of the technology. This helped many African Leaders move the implementation process forward due to the level of awareness or pressure from civil society and private sector.

References

1. L. Adam, N. Butcher, F.F. Tusubira, C. Sibthorpe, Transformation-ready: the strategic application of information and communication technologies in Africa education sector study, 2011
2. E. Adomi, *Application of Information and Communication Technologies (ICTs) in Nigerian High Schools* (Nigerian Library Association, Delta State Chapter AGM, Warri, 2006)
3. L.O. Aina, Coping with the challenges of library and information delivery services: the need for institutionalized professional development, in *Nigerian Library Association Conference Proceedings*, 2004, p. 4
4. G.O. Ajayi, African Information Society Initiative (AISI) An Action Framework to Build Africa's Information and Communication Infrastructure. Economic Commission for Africa, 1995
5. G.A. Alabi, Empowering socio-economic development, utilizing information technology: a critical examination of the social, economic, technical and policy issues in Nigeria, 2004
6. A. Aman, B. Nicholson, *The Process of Offshore Development: Preliminary Studies of UK Companies in Malaysia* (Kluwer, Boston, 2003)
7. T.N. Andrew, A Systems Thinking Approach to the Planning of Rural Telecommunications Infrastructure. Ph.D. thesis (University of Natal, South Africa, 2001)
8. K. Annan, Message from United Nations Secretary-general Kofi Annan, 2003
9. APC (Association for Progressive Communications). 2000. Using the internet strategically
10. J. Aswalap, Information society development in Thailand: information and communication technology perspectives, 2005
11. D. Banard, T. Vonk, Information and Communication Technologies in African Local Governments. EDITOSIA – Electronic Distance Training on Sustainability in African Local Governments. ICLEI European Secretariat GmbH. Freiburg, Germany, 2003
12. A. Bhavnani, R. Won-Wai Chiu, S. Janakiram, P. Silarszky, (TTL) The role of mobile phones in sustainable rural poverty reduction, 2008
13. J. Braa, C. Hedberg, The struggle for district-based health information systems in South Africa. Inf. Soc. **18**(2), 113–127 (2002)
14. A.E. Brown, G.G. Grant, Highlighting the duality of the ICT and development research agenda. Inf. Technol. Dev. **16**(2), 96–111 (2010)
15. M. Castells, *The Rise of the Network Society* (Blackwell, Boston, 1996)
16. M. Castells, Globalization in the network. Prometheus J. **4**, 109–123 (2000)
17. R. Chambers, Paradigms, poverty and adaptive pluralism. IDS Working Papers 344, 2010, pp. 01–57
18. B. Chilundo, S. Sahay, J. Sundby, Analysing the quality of the HIV/AIDS case-detection and case-reporting systems in Mozambique. Afr. J. AIDS Res. (in press)
19. J. Chisenga, Global Information and libraries in sub-Sahara Africa. Libr. Manag. **21**(4), 178–187 (2000)
20. L. Colin, W. David, Perspectives on Distance Education: Telecentres: Case studies and key issues (Management|Operations|Applications|Evaluation), 2001
21. P. Constable, Afghan women take radio liberties: tiny station transmits messages of support to a largely illiterate female populace. Washington Post Foreign Service, November 3, 2003; Page A10 (2003)
22. L. Cuban, Responding to the Twin Challenges: Teachers, Teaching, and Technology, 1986
23. M. D'Mello, Thinking local, acting global: an analysis of identity related issues in a GSO in India, 2003
24. T.R. Enakrire, G.O. Onyenania, Causes Inhibiting the growth or development of information transfer in Africa: a contextual treatment. Library Hi Tech News **24**(4), 20–28 (2007)
25. G. Farrell, S. Isaacs, M. Trucano, Survey of ICT and education in Africa: a summary report based on vi 53 country surveys, 2007
26. G. Farrell, S. Isaacs, M. Trucano, *Survey of ICT and Education in Africa: 53 Country Reports,* vol. 2 (infoDev/World Bank, Washington, 2007)

27. M. Faye, Developing national information and communication infrastructure (NICI): policies and plans in Africa. Paper presented during the Nigeria NICI workshop, Abuja, Nigeria, 2000
28. R. Gerster, S. Zimmermann, Information & Communications Technologies (ICTs) & Poverty Reduction in Sub Saharan Africa. CH – 8805, Richterswil, Switzerland (2003)
29. N. Hafkin, N. Taggart, *Gender, Information Technology and Developing Countries: An Analytic Study* (WID Office USAID, Washington, DC, 2001)
30. R. Harris, Information and communication technologies for poverty alleviation, the United Nations Development Programme's Asia-Pacific Development Information Programme (UNDP-APDIP), 2004
31. ICT, MDGs (2003) A World Bank Group Perspective, December 2003
32. P. Ifinedo, Measuring Africa's e-readiness in the global networked economy: a nine-country data analysis. Int. J. Educ. Dev. Inf. Commun. Technol. (IJEDICT) **1**(1), 53–71 (2005)
33. Innovation (2004) Economist, pp. 67–68
34. International Telecommunication Union (ITU), *World Telecommunication Indicators Database* (International Telecommunications Union (ITU), Geneva, 2008)
35. International Telecommunication Union (ITU), *Information Society Statistical Profiles 2009 Africa* (International Telecommunications Union (ITU), Geneva, 2009)
36. ISRDS, The Integrated Sustainable Rural Development Strategy. Republic of South Africa Government Documents, 2000
37. ITD, *Information Technology for Development*, vol. 12(1) (Wiley Periodicals Inc., New York, 2006), pp. 7–24 C_2005
38. ITU, *Measuring the Information Society. The ICT Development Index* (ITU, Geneva, 2009)
39. ITU-D, 2010, http://www.itu.int/ITU-D/ict/statistics/at_glance/KeyTelecom2010.html
40. IUSA, Internet usage statistics for Africa, 2011
41. M. Korpela, R. Montealegre, A. Poulymenakou, in *Proceedings of the International Federation of Information Processing, IFIP 9.4 and 8.2 Joint Conference on Organizational Information Systems in the Context of Globalization* (Kluwer, Dordrecht, 2011), pp. 201–216
42. LA, 2011, http://www.lightingafrica.org/component/k2/item/22.html?layout=item
43. E. Macome, On the implementation of an information system in the Mozambican context: the EDM case, 2003
44. S. Madon, The internet and socio-economic development: exploring the interaction. Inf. Technol. People **13**(2), 85–101 (2000)
45. A. Maiye, K. McGrath, in *ICTs and Sustainable Development: A Capability Perspective, AMCIS 2010 Proceedings*, Paper 541, 2010
46. M. Metcalfe, C. Joham, The "ear" and "eye" digital divide, eds. by M. Korpela, R. Montealegre, A. Poulymenakou, *Proceedings of the International Federation of Information Processing*, IFIP 9.4 and 8.2 Joint Conference on Organizational Information Systems in the Context of Globalization (Kluwer, Dordrecht, pp. 419–434, 2003)
47. E. Mosse, S. Sahay, Counter networks, communication and health information systems: a case study from Mozambique, 2003
48. S. Mutula, IT diffusion in sub-Sahara Africa: implications for developing and managing digital libraries. New Library World **102**(1202/1203), 281–289 (2004)
49. National Policy on Telecommunications (NPT), 1999
50. Neil Butcher Associates (2010) ICT, Education, Development, and the Knowledge Society – Thematic paper prepared for African Leaders in ICT: Building Leadership Capacities for ICT and Knowledge Societies in Africa; and World Bank. ICT and Education-Key issues
51. A. Ngwenyama et al., Is there a relationship between ICT, health, education and development? An empirical analysis of five West African countries from 1997–2003. Electron. J. Inf. Syst Dev. Countries **23**(5), 1–11 (2006)
52. L.A. Ogunsola, Information and communication technologies and the effects of globalization: twenty-first century "digital slavery" for developing countries – myth or reality? Electron. J. Acad. Special Librariansh. **6**(1–2) (2005)
53. J. Okpaku, Regional workshop on building e-governance capacity in African countries background paper on ICT-for-development in Africa for the United Nations ICT task force, 2002

54. A. Okunoye, H. Karsten, Global access to knowledge: findings from academic research organisations in sub-Saharan Africa. Inf. Technol. People **16**(3), 353–373 (2003)
55. W.J. Orlikowski, J.J. Baroudi, Studying information technology in organizations: research approaches and assumptions. Inf. Syst. Res. **2**(1), 1–28 (1991)
56. T. Oyedemi, Universal access wheel: towards achieving universal access to ICT in Africa. Southern Afr. J. Inf. Commun. **5**, 90–107 (2003)
57. H.H. Rolland, E. Monteiro, Balancing the local and the global in infrastructural information systems. Inf. Soc. **18**(2), 87–100 (2002)
58. E.H. Sayed, C. Westrup, Egypt and ICTs bring national initiatives, global actors and local companies together. Inf. Technol. People **16**(1), 93–110 (2003)
59. L. Silva, E.B. Figueroa, Institutional intervention and the expansion of ICTs in Latin America: the case of Chile. Inf. Technol. People **15**(1), 8–25 (2002)
60. The Economist, Halfway there: how to promote the spread of mobile phones among the world's poorest. GSMA, 2008
61. Trucano Michael, *Knowledge Maps: ICTs in Education*, (infoDev/World Bank, Washington, 2005)
62. UNCTAD, Information Economy Report 2010: ICTs, Enterprises and Poverty Alleviation, Technical report, in *United Nations Conference on Trade and Development (UNCTAD)*, 2010
63. UNDP, Human Development Report: Making Technologies Work for Human Development (Washington, 2001)
64. J. Von Braun, ICT for the poor at large scale: innovative connections to markets and services, in *ICT for the Next Five Billion People: Information and Communication for Sustainable Development*, ed. by A. Picot, J. Lorenz (Springer, Heidelberg, 2010), pp. 3–14. Chapter 2
65. Vukanikids-DTI, An analysis of the ICT skills audit for the dti and ISETT SETA. Department of Trade and Industry (DTI) –South Africa, 2005
66. Waema TM (2002) ICT Human Resource Development in Africa: Challenges and Strategies
67. Waema TM (2002) ICT Human Resource Development in Africa: Challenges and Strategies
68. G. Walsham, Knowledge management: the benefits and limitations of computer systems. Eur. Manage. J. **19**(6), 599–608 (2001)
69. G. Walsham, Cross-cultural software production and use: a structurational analysis. MIS Quart. **26**(4), 359–380 (2002)
70. L. Waverman, M. Meschi, M. Fuss, The impact of telecoms on economic growth in developing countries, in *Africa: The Impact of Mobile Phones*. The Vodafone Policy Paper Series, vol. 2 (Vodafone, Newbury, 2005)
71. World Bank (2002a)
72. World Bank (2003b)
73. World Bank, From Privilege to Competition. World Bank, 2009
74. F.A. Yiek, Towards alternatives in higher education: the benefits and challenges of e-learning in Africa. CODESRIA Bull. **3 & 4**, 73–75 (2005)

Part III
On Innovative Technologies: Focus on Embedded Systems

The Co-design of System-on-Chip in Africa: Spotlights on Industry, Research Efforts, Trends, and Challenges

Manel Ammar, Mouna Baklouti, and Mohamed Abid

Abstract On a worldwide basis, Europe, United States of America (USA), and Asian countries continue to reinforce their leading position and intensify their market share of the global worldwide sales of semiconductors and electronic design automation (EDA) tools. In the emerging field of system-on-chip (SoC) design, research efforts targeting co-design methodologies and tools in these countries follow the same direction. During the last decade, African research efforts in the SoC design domain continue to grow and follow the worldwide evolution and trends. This chapter provides an analysis of the major technological advances in the semiconductor and EDA industries with a special focus on the situation of Africa among this progress. This chapter will also detail the two ages of SoC design research efforts in Africa: the age of language-based design flows and the age of model-based co-design flows. This in-depth paper provides an analysis of the SoC design process including approaches, trends, challenges, and perspectives making Africa in the center of interest.

1 Introduction

The exponential rise of tiny electronic devices number in numerous products is common knowledge. These devices, named integrated circuits (IC), contain millions of transistors gathered into a small surface. Designing an IC becomes easier with the emergence of computer-based tools. First, IC designers use Computer Aided Design (CAD) tools for the placement and route tasks. Then, with the evolution of the semiconductor industry, other tools appear including Computer Aided Engineering (CAE) and Computer Aided Manufacturing (CAM) tools. Electronic Design Automation (EDA) is a generic term for all these tools. Because electronic circuit design is a less tolerant mean of design, IC designers need more powerful

M. Ammar (✉) • M. Baklouti • M. Abid
CES Laboratory, National Engineering School of Sfax, Sfax, Tunisia
e-mail: manel.ammar@ceslab.org; ammar_manel@yahoo.fr; mouna.baklouti@enis.rnu.tn; mohamed.abid@enis.rnu.tn

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_9

tools to deal with long design cycles and hard constraints. For this reason, system engineers, logic designers, and layout designers use EDA tools to facilitate and automate IC design.

An IC that integrates software and hardware components into a single chip is known as System-on-Chip (SoC). While the design of an IC is a complex task with the evolution of the semiconductor industry that permits to integrate a huge number of transistors in a single chip, the design of SoCs is harder when software and hardware teams must work together concurrently to deliver a reliable product with respect to the time-to-market and cost constraints. This incredible evolution in the IC and SoC domains could not be achieved by hiring an exponentially growing number of design engineers. But, making the productivity of SoC designers stronger by adopting new design methodologies supported by EDA tools allows to strengthen the SoC designers' efficiency. These tools are based on the co-design methodology that gathers the design of hardware and software in a unified flow.

For many years, the largest electronic companies that consist of semiconductor companies and EDA companies, concentrated in the USA, Europe, and Asia had developed complex SoCs and powerful EDA tools that support the design of high-performance systems and circuits. At the same time, Research & Development (R&D) teams had tried to help these companies elaborating new methodologies and defining other levels of abstraction and models of specification. Putting African companies and African academic research teams in the spotlight and focusing on their efforts done in the SoC co-design domain is necessary when the worldwide industry of semiconductors and EDA tools is growing exponentially. It is the aim of this chapter to study the evolution of the semiconductor and EDA industries in Africa and to study the academic research efforts that define novel SoC design approaches.

This chapter is structured into seven sections. Section 2 focuses on the evolution of the semiconductor and the EDA industries in the world and discusses the place of the African industry regarding this evolution. Concepts related to SoC design approaches are explained in Sect. 3. This section describes two categories of SoC design methodologies: the classic approach and the co-design approach with special focus on challenges and new trends. Section 4 describes several African academic research efforts with a special focus on the proposed design flows. These efforts are classified into two ages. Section 5 describes some worldwide related works. A detailed evaluation of African efforts is given in Sect. 6.

2 A Focus on Embedded Systems Industry

Semiconductor devices provide solutions that help designers to enhance energy consumption and performance of everyday technology applications. Cars, for example, are an army of semiconductor-powered electronics. Complex electronic systems are semiconductor chips realized using SoC technology, which can integrate heterogeneous digital components on a single chip. The EDA industry is developing

Table 1 Big semiconductor companies in Africa

Company	Rank	Country of origin	Locations in Africa
Qualcomm	3	United States	Nigeria, South Africa, and Kenya
STMicroelectronics	8	France and Italy	Tunisia and Morocco

tools to help designers creating complex SoCs. This section tries to find the place of the African industry in the worldwide electronic industry. While the first part of this section discusses the evolution of the semiconductor industry, the last part focuses on the EDA industry.

2.1 Evolution of the Semi-conductor Industry

Gordon Moore assumed in 1965 that the number of transistors in a single chip would double every 2 years [36]. Following this historical trend, named Moore's Law, the density of transistors doubled every 1.96 years between 1971 and 2001. As a result, the relative manufacturing cost per function has become less expensive, facilitating the production of more complex circuits on a single chip.

The semiconductor technology has been progressing rapidly for more than 40 years following this law. In 2000, the semiconductor industry achieved $250 billion by decreasing the cost per elementary function. This enables the emergence of new systems such as telecommunication systems, radios, PCs, and TVs.

At the Intel Developer Forum in September 2007, Gordon Moore predicted that his famous law would no longer be valid in 10–15 years. For this reason, two new directions labeled "More Moore" and "More-than-Moore" have been emerged to replace Moore's law in the International Technology Roadmap for Semiconductors (ITRS) reports [33]. The "More Moore" trend promotes the miniaturization of the digital functions using new transistor concepts and incorporating into devices new materials. The More-than-Moore direction focuses on the non-digital functional diversification of devices.

Semiconductor manufacturers take advantage of these fast changes and continue to compete and provide the latest technologies, even though costs are increasing. According to the IHS iSuppli Industrial Electronics Market Tracker Report, the semiconductor industry is leaded by Japan, Korea, Taiwan, USA, Singapore, and Europe with total revenue of $323.2 billion in 2012 [50]. Intel Corporation led the top 10 suppliers of semiconductors according to the same report. An important question is where African countries in the worldwide classification of semiconductor manufacturers are.

To make fast progress in semiconductor technologies and avoid cost increase, some manufacturers have a substantial labor force in lower cost locations such as African countries. Looking at the Table 1, STMicroelectronics owns two design centers in Africa. The first is in Morocco with 160 employees and the second

in Tunisia with 300 employees. Qualcomm, owning the third place in the top 10 ranking of semiconductor companies, has three offices in Africa: the first in South Africa, the second in Nigeria, and the third in Kenya. We can notice that African countries are regarded as a solution for reducing the design cost. But none of these countries appears to be a leader in the semiconductor industry. In fact, these countries remain consumers of the semiconductor technology.

2.2 Expansion of EDA Tools

Over time, innovation is being a big challenge for the electronic industry. In fact, the whole industry is moving up from simulation and place and route to high level design.

In the 1960s, five fundamental tendencies have emerged: circuit simulation, logic simulation and testing, Metal Oxide Semiconductor (MOS) timing simulation, wire routing, and regular arrays. During this period, CAD was regarded as a strategic value to the electronic industry. International Business Machines (IBM) Corporation considered it important enough to deserve a considerable investment in both funding and resources. Other large companies were also interested in CAD tools among them we can mention Intel. Its first design environment was mainly created to supply the needs of memory chips. During the 1970s, layouts were too difficult to develop and maintain. For this reason, CAD tools were created to generate and check mask layouts. Between 1979 and 1993, the electronic design field exploded in all its aspects. Verification and testing, layout, logic synthesis, hardware description languages (HDL), hardware acceleration, parallel computers, and system level design have been emerged as innovative contributions. In the late 1980s, a big deal of interest in parallel architectures was greatly present in the computer design community. During this phase, EDA was in the spotlight and similarly engaged in both computer science and electrical fields.

EDA for electronics becomes more and more crucial with the continuous scaling of semiconductor technology. The year of 1981 denotes the establishment of EDA as an industry. For many years, numerous electronic companies, such as Synopsys, Cadence, Mentor Graphics, and Zuken Inc, had leaded the EDA market. Similar to the semiconductor pioneers, these companies are founded in USA and Japan. In response to the high cost of the EDA tools development, companies have resorted to the lower-cost locations such as Asia. Mentor Graphics, ranked the third in the EDA industry in 2013, takes place in Egypt. This powerful company opened its doors to the north of Africa for experience exchanging and common EDA tools development. But, other African countries lack big EDA companies' presence.

After focusing on the semiconductor industry and before describing academic efforts in the SoC design field, an overview of the SoC design concepts, challenges, trends, and categories of approaches will be outlined in the next section.

3 Design of Embedded Systems: From the Classical Approach to the Co-design Approach

It is the purpose of this chapter to shed some light on the recent practices of the SoC design approach in Africa. Our goal is to outline several design approaches, to show the scientific contribution of each approach, and to discuss differences and similarities. For this reason, we describe first the evolution of the design flows of integrated circuits distinguishing their characteristics that are useful to classify the different approaches.

3.1 The Classical Approach: Separation of Concerns and Low-levels of Abstraction

Separation of concerns and low-levels of abstraction are the most important characteristics of the classic design methodologies. These characteristics will be outlined in this subsection.

Separation of Concerns

With the appearance of integrated circuits, hardware is regarded as an assembly of electronic components including processing units, memories, communication systems, and other peripherals. Software is considered a program stored in the memory and executed by the processor. At that time, both hardware and software designs progress separately and partitioning decisions are fixed at an early stage in the development cycle. The choice to implement an operation in hardware or software is done manually based on the design team experience. Other design constraints are taken into consideration in this phase including cost, performance, technological limits, and others. Designers are separated into different teams; Application-Specific Integrated Circuit (ASIC) designers represent the hardware team and software designers stand for the software team. While the hardware team has a hardware-centric view of the system design problem, the software team has a software-centric view; the fact that creates a large boundary between hardware and software design progress.

The emergence of field-programmable gate array (FPGA) systems has destroyed the separation between hardware and software parts of the system. Using the classic design flow, the hardware part of the system, which remain unmodified, can be configured only at manufacturing time. With the appearance of FPGA, the designer can configure the circuit after manufacturing. This flexibility promotes the design of complex applications and their integration into these circuits. Managing the design of these systems becomes a challenge for both system and semiconductor industries

that used a purely software or hardware strategy. The software and hardware teams must think of embedded system design holistically, creating a strategy that design software and hardware concurrently and not independently.

Low-levels of Abstraction

In the history of design flows, changes in design productivity were always related to raising the level of abstraction in design entry. In the 1970s, the highest level of abstraction was a transistor schematic. Ten years later, with the appearance of CAD tools in 1980, design entry has moved up from transistors to gates. Then, with the appearance of HDL by 1990, and the evolution of CAD tools, other levels of abstraction were proposed including the Register-Transfer Level (RTL) and the behavioral level. VHSIC Hardware Description Language (VHDL) and Verilog are used to describe the hardware part of the system at the RTL or behavioral level. RTL synthesis allows deriving RTL description of the system from the behavioral description. Then, logic synthesis is applied from the RTL to the logic-gate layer of abstraction. When applications become more complex and chips density continues to expand exponentially, higher levels of abstraction need to be created.

3.2 The Co-design Approach: Toward New Trends in the Design of SoCs

Designers of high performance embedded systems are facing many design challenges. This subsection discusses some of these critical challenges and introduces possible solutions.

SoC Design Challenges

Nowadays, there is a big development in the domain of SoC. Reducing the physical surface area, reducing the consumed energy, and diminishing the overall fabrication cost are the major advantages of SoC. Embedded systems, currently, are generally dedicated to high performance computing (HPC) where many processor cores are used to gain in performance.

The rapid evolution of applications complexity and the increase of integrated technologies capacities are evolving side by side. Unfortunately, design flows and development tools have not been able to follow such developments. The result is a gap in productivity intensified by economic laws. Indeed, methods and tools used in the design and verification phases are not suitable to manage the increasing complexity of software application and hardware architecture. ITRS Semiconductor roadmap [24] shows that the demand for software is doubling every

10 months. The same report noted that the SoC integration capabilities double every 36 months and software productivity also doubles every 5 years. Hence theoretical and experimental SoC integration ability continues to grow.

Need of Well Structured Methodologies: The Co-design Approach

With the density of current and future design requirements, design productivity for wide-ranging SoC is a notable challenge. The semiconductor industry and its customers would benefit by using CAD tools. According to the ITRS report [23], using CAD tools can reduce the design cost of embedded systems by 40 % in 2013. The long development time can be also reduced by 40 % in the same year. The expansion of CAD tools raises the interest in hardware/software co-design. Co-design [20], an approach for parallel hardware and software design, is considered a mean to accelerate the design cycle and reduce the integration time. It is a multi-step design methodology, which takes the system specification as input, then moves up from Design Space Exploration (DSE) to partitioning and scheduling, then synthesis, and finally placement and route.

Need of Abstract Models

The raising complexity of embedded systems creates the need of intensive specification task. Added to that, quality is naturally required in systems design to deliver competitive products. This requirement must be satisfied from the first stage of the design flow by making complete specifications. In this scenario, the specification task can be defined as a delicate design phase that is very demanding in time and cost. These dilemmas can be solved, thanks to the design abstraction that certainly facilitates the design task and favors an efficient design reuse. This abstraction is expressed via models that capture the specification of the system from different viewpoints. These models serve as an entry point of the design flow and they are refined during the different phases of an embedded system design process.

Need of Platform-Based Design

The Platform-Based Design (PBD) methodology [27] exploits the design reuse approach during all phases of design process. A platform is defined as a library of components having the same abstraction level. These components can be brought together to produce a design at that level of abstraction.

This approach enhances the development of parametric architectures. So, it is a powerful concept for scalability as it offers the possibility to design several applications and deploy them on these architectures on the one hand, and it guarantees performance of reuse on the other hand.

Need of SoC Design Verification

Integrating hardware and software components remains the most delicate step in embedded systems design. This step must guarantee that the embedded system executes the desired tasks perfectly on the one hand, and meets its design requirements on the other hand. Both verification and validation enable the successful integration of hardware and software parts of the embedded system. While the verification aims to prove the correctness of the system, the validation demonstrates whether the system constraints are satisfied or not.

The ever increasing complexity of modern embedded systems and the downgraded lifespan of current semiconductor products have made the SoC verification and validation tasks very difficult. For these reasons, researchers have tried to place these tasks early in the design process. Charfi et al. [11] integrate a validation tool, named GaspardValidation, in the Gaspard2 [18] framework for the validation of SoC design. Models of the embedded system are verified from the higher level of abstraction, the Unified Modeling Language (UML) models level. The Object Constraint Language (OCL) [45] enables to specify the properties to verify and the proposed validation tool checks the correctness of the defined models. While the OCL language enables the non-formal verification of the design from the specification level, formal languages have been also emerged to provide a formal verification of the design. An example of using a formal language for the verification purpose can be found in [22]. Authors define a methodology for the automatic translation of SystemC descriptions into SIGNAL models for verification and validation purposes. Structure and behavior of the system are extracted and translated into formal descriptions. Then, these generated models are used to verify the embedded system.

4 Two Ages of Co-design Research Efforts in Africa

African research teams in Tunisia, Algeria, and Egypt have concentrated their efforts in defining new methodologies that facilitate the design of SoC during the last decade. While early approaches support language-based design entries targeting simple models of architecture, recently developed flows are model-based approaches that target complex models of SoC.

4.1 The Age of Language-Based Design Flows

In this period, first design efforts were concentrated on defining a methodology for mixed hardware and software design [1]. These efforts define the co-design approach as a new method that merges the software engineering and the hardware engineering to be able to specify, design, and validate the software and hardware parts of an embedded system simultaneously. Creating a unified environment for the

description, partitioning, synthesis, and verification of embedded systems where the hardware and software parts can be interfaced throughout the design process is the major goal of this work. The key points of the proposed methodology concern:

- Defining unified software and hardware models: In this work, the author identifies two approaches for the specification of an embedded system in a co-design flow. The first approach is based on one language for the specification of hardware and software parts. The second approach, supported in this work, specifies the two parts of the system using both software and HDL and tries to integrate them in a single co-design environment. While software models describe the logical structure and the behavior of the system using the C language, the architecture describes how this system is made in terms of physical structure and implementation. These hardware implementations are written in VHDL in the behavioral level of abstraction.
- Implementing a flexible and modular architecture: The author proposes a generic model of the architecture. In such a model, the system is organized as a set of software modules, hardware modules, and communication modules.
- Using a rapid DSE approach: The proposed approach considers a mixed representation of the system allowing to ensure the co-simulation and co-synthesis into a unified environment as described in Fig. 1a. This allows exploring the solution space by co-simulation and generation of multiple implementations on a target platform from a single abstract model. It is an iterative exploration approach that takes into account the performance of the target architecture. Performance is expressed using different metrics including execution time and implementation costs. Existing synthesis and simulation tools are combined with implemented ones for the estimation and evaluation of feasible solutions.
- Developing a strategy for the refinement of the design from high-level descriptions: In a typical synthesis process, embedded system design is regarded as a progressive top-down flow involving several stages of refinement to achieve the layout generation. Independently from the considered level of specification, synthesis can move from one level of abstraction to another given at a lower level, while preserving the functionality of the circuit. Automation in part or in the whole of this process facilitates the design of embedded systems. The proposed approach presents a typical synthesis flow that takes VHDL files as design entry and generates the netlist through a number of synthesis steps including behavioral synthesis, RTL synthesis, and logic synthesis. A simulation of the system is done after each synthesis step. Finally, three-step rapid prototyping process takes place including hardware prototyping, software prototyping, and system prototyping.

Automating software/hardware co-design was the focus of another research paper [17]. Authors propose a purely software-based design flow where the embedded software is written in C (Fig. 1b). Starting with a software rich description aims to make the exploration faster, to use existing co-verification tools for the C code, and to facilitate the integration of a profiler in the co-design flow for early estimations. Based on the gathered profiling data, the flow decides if it is necessary to convert several software functions to hardware. Functions needed to be transformed into hardware components are generated in HDL files.

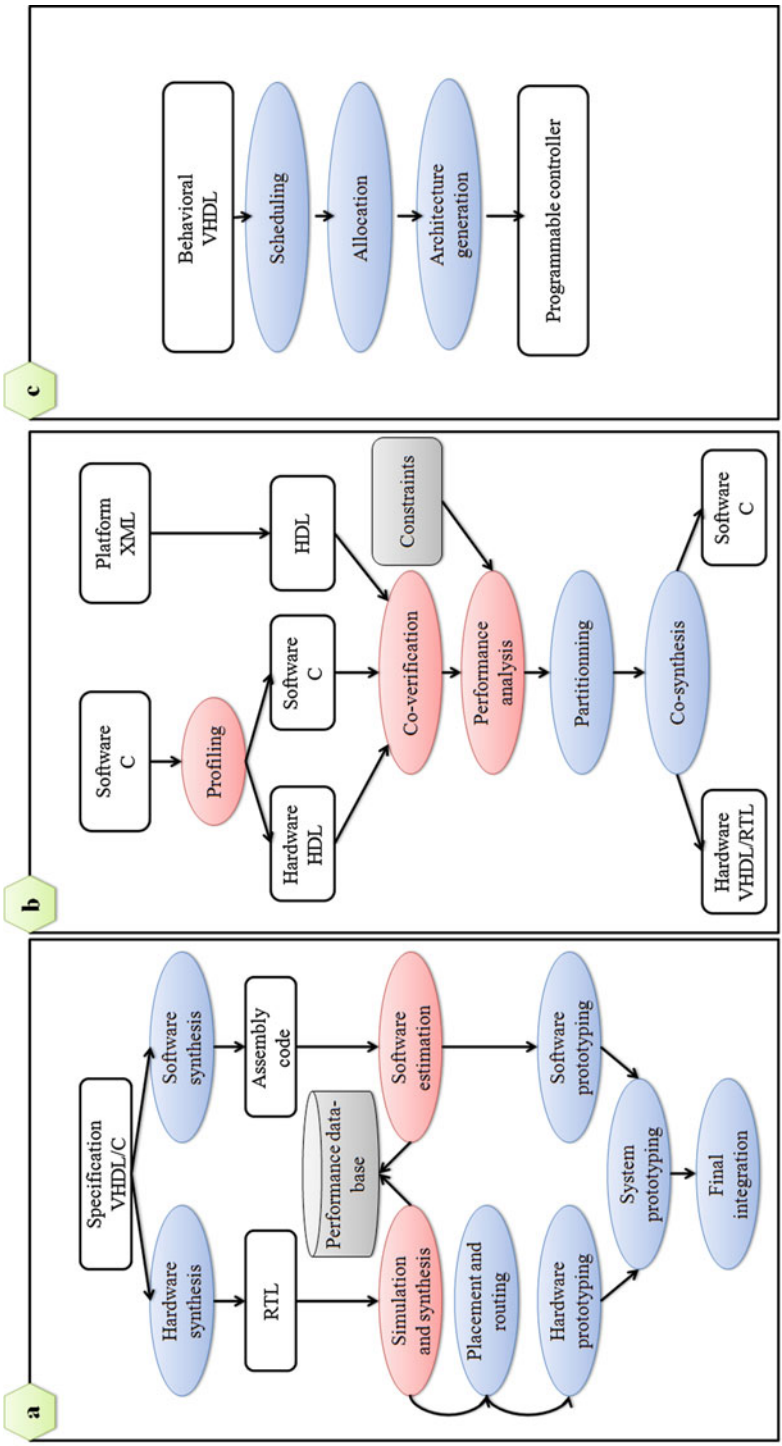

Fig. 1 Language-based design flows

This work supports the PBD approach. The platform specification is captured graphically using a PBD tool. Then, an XML description of the platform is generated. The co-specification includes, apart from the software specification and the platform specification, system criteria modeling and verification (using co-simulation).

The co-verification of the system is carried out after the integration of the platform description (XML), the software specification (C), and the hardware components files (HDL). During the integration, the XML platform description is transformed into HDL files. At this stage, C and HDL files can be easily compiled and verified using a unified co-verification tool [35]. After that, obtained performance of the system is compared with the specified criteria from the co-specification phase. If the obtained performance is sufficient, the flow will be terminated and the system delivered for implementation. Otherwise, a partitioning phase will be executed for further optimization. The last phase in this co-design flow is the co-synthesis. This step includes communication synthesis and specification refinement. While the C code of the software remains the same in the specification refinement stage, the hardware synthesis generates VHDL code at the RTL abstraction level.

Another research work, presented in [6], denotes a co-design flow dedicated to the automatic generation of microcoded for Application Specific Instruction-set Processor (ASIP) controllers. This work (Fig. 1c) extends an existing architectural synthesis system, the AMICAL tool, to support the design of reprogrammable microcoded controllers.

AMICAL [28] helps the designer to automatically generate an architecture that tunes for different applications using control-flow dominated machines as design entry. This high-level synthesis tool allows the scheduling, the allocation, and the architecture generation tasks combining manual and automatic synthesis. AMICAL targets only hardwired controllers and do not support programmable ones. This work is concentrated on the transformations added to the AMICAL framework targeting ASIP controllers [7]. From a behavioral description of the system using the VHDL language, the architecture code is generated after allocation and scheduling.

4.2 The Age of Model-Based Co-design Flows

Recently, enormous efforts have been made to elaborate design methods and tools that solve embedded system design issues described in Sect. 3. In fact, reaching higher levels of abstraction is the major focus of SoC designers with the rising of SoC density and the expansion of high-performance applications. Since refinement keeps the properties of the higher abstraction [47], the abstraction level of the initial specification defines the complexity of the design approach and the set of solutions to be considered. For that reason, we classify the design approaches, taking the abstraction level of the specification as base, into two categories: graph-based approaches and UML-based approaches.

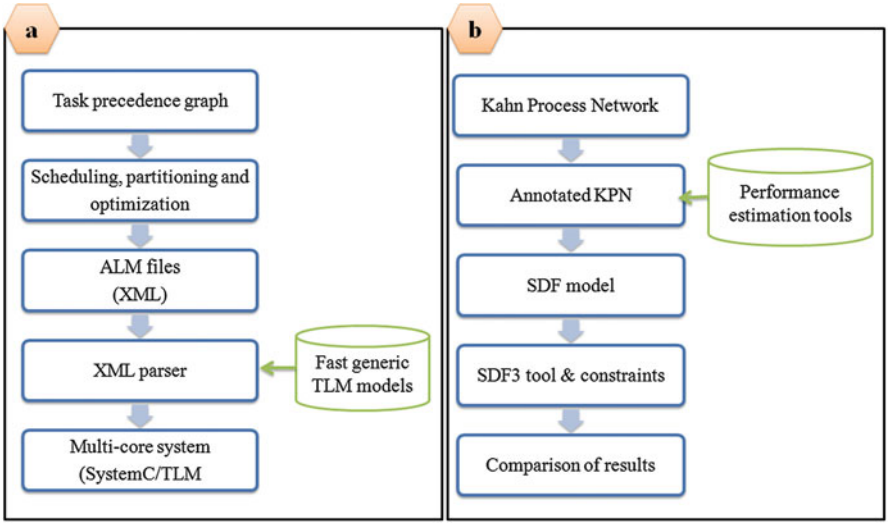

Fig. 2 Graph-based design flows

Graph-Based Approaches

In a graph-based approach, an application is specified as a set of tasks, or actors, connected to each other through channels representing data flows and dependencies. There are numerous graph models that have been proposed helping to specify modern embedded systems.

A research group from Mentor Graphics Corporation and Ain Shams University in Cairo, Egypt, takes advantage from a graph-based approach for the architecture exploration of MPSoCs [51, 54]. The design entry of the proposed flow is a software application specified using a Task Precedence Graph (TPG) (Fig. 2a). A TPG is composed of nodes representing the set of tasks and directed edges denoting the execution dependencies and the amount of communication. This Model of Computation (MoC) is appropriate for static scheduling of parallel algorithms on MPSoCs [30].

The proposed flow automatically generates diverse architectures having different number of processor cores and busses. In addition, it provides mappings of tasks to the generated set of processors and channels to the generated busses.

The design flow starts with an annotated graph. Each task is annotated with its computation time and channels between tasks are enriched with communication time. Added to that, energy consumption for computations and communications are identified in advance.

A scheduling algorithm is then executed; this algorithm is detailed in [55]. The algorithm is responsible of solving the allocation and scheduling problems. An optimal schedule on a multiprocessor system having a reduced processor number is then generated [56]. Conflicts in the communication channel are also resolved using

a graph coloring algorithm. A hardware/software partitioning technique is used to reduce the execution time of the application converting slow software tasks of the graph into hardware components and accomplishing channel mapping.

The task mapping, the channel mapping, and the HW/SW partitions are gathered to provide several Architecture Level Models (ALM) files. Each ALM file denotes an architecture structure specified by the number of cores, the number of busses, the task-to-core mapping, and the channel-to-bus mapping. An XML parser is used in the last phase of the flow to generate the multi-core system, described in SystemC at the Transaction-Level Modeling (TLM), from ALM files. This phase is guided by a library of fast generic TLM models.

Another graph-based co-design flow for MPSoC is presented in [53]. Instead of using the previously described TPG MoC, authors adopt a Kahn Process Network (KPN)-based model [25] for the specification of the system. The proposed approach (Fig. 2b) facilitates the performance estimation of migration software tasks to hardware components using four steps:

- Annotating the KPN model: KPN model is widely used for specifying distributed systems and parallel programming. It is a dataflow network with a set of nodes presenting concurrent processes and unidirectional First In, First Out (FIFO) channels. Estimating the performance of MPSoC system needs information related to time and data size. For this reason, the KPN model of application in this approach is annotated with the execution time of each task and the size of data exchanged between these tasks. These metrics are obtained using performance estimation tools.
- Transforming KPN to SDF: SDF [32] is one of the most fundamental data flow models having a wide range of variants. This actor-based model is widely used in signal-processing applications. Actors and channels, the main components of the dataflow, give semantics to facilitate performance analysis, scheduling, DSE, and other design process for embedded systems. For this reason, the annotated KPN parallel model is transformed into an SDF model. The developed transformation depends on the application, as well as the hardware architecture. For these reasons, the generated model is a timed SDF graph enriched with numerous parameters (type of communication system, memory structure, etc.)
- Executing a synthesis under constraint: If the SDF model is generated, another step will be executed for performance estimation, the synthesis. This step is based on the SDF3 tool that has been reconfigured to realize the synthesis under constraints. This tool can be used for estimating the performance of the applications on Network-on-Chip (NoC)–MPSoC systems. This estimation is based on a static analysis of the SDF graph. Functional and non-functional constraints must be taken into account in this approach. For this reason, a formalization of the restrictions is first accomplished, then, an extension of the SDF3 tool is realized.
- Comparing results: This step is dedicated to analyze estimated performance gains with SDF graphs and those measured using the traditional way, the simulation.

The Sesame framework [13] offers capabilities for the modeling and simulation of heterogeneous embedded multimedia systems. It is a system-level co-design flow based on three layers for the system specification. While the functional behavior of the application is described using KPN in the application modeling layer, the architecture resources and their performance constraints are defined in the Pearl simulation language or in SystemC in the architecture modeling layer. The mapping layer performs both Hw/Sw mapping and scheduling. A trace-driven simulation is executed to evaluate the system performances.

A research effort in Tunisia [58] takes advantage of the previously described tool to evaluate the performance of a H. 264/AVC encoder for multiprocessor architectures. First, a parallel implementation of the encoder is obtained, thanks to a high-level independent target-architecture parallelization methodology developed by the same author and described in [29]. The obtained model is optimized and specified in KPN. Using components of the architecture library permits to obtain four platform models varying the processors number. These models capture the architecture layer, while the optimized KPN model presents the application layer. Performance measurements of the H. 264/AVC encoder are finally estimated using a flexible co-design approach that takes advantage of existing tools to support new applications.

UML-Based Approaches

To define model-based specification, we need to use a formal language such as the UML [41]. This formalism offers attractive graphical specification concepts that help the designer to specify his system clearly. It is true that UML is a general language but its extensibility permits to address domain-specific problems. This is mainly done via the notion of profile. A profile is then a collection of extensions and restrictions that describe a particular domain. This extension mechanism is generally expressed using stereotypes, tag definitions, and constraints. Stereotypes enable designers to extend UML to make new model elements that derive from existing ones [42]. These model elements have specific properties that are appropriate for a particular domain named tag definitions.

Authors in [10] present a UML profile and a co-design methodology for Worst-Case Execution Time (WCET) estimation of embedded systems (Fig. 3a). Sequence diagrams of UML are used as a high-level design entry offering a pure sequential model of the application. Interaction between objects, data and control dependencies, loops and conditions, and finally temporal constraints (WCET) are well defined using the sequence diagram. The WCET of each object method in the sequence diagram is obtained using profiling or introduced manually. Starting from a sequential object paradigm, diagrams are converted into a hierarchical task graph. This intermediate model is composed of a set of communicating tasks and communication channels. Tasks expose task parallelism, data parallelism, pipelining, and hierarchy. The task graph is enriched with stereotypes of the proposed profile [8].

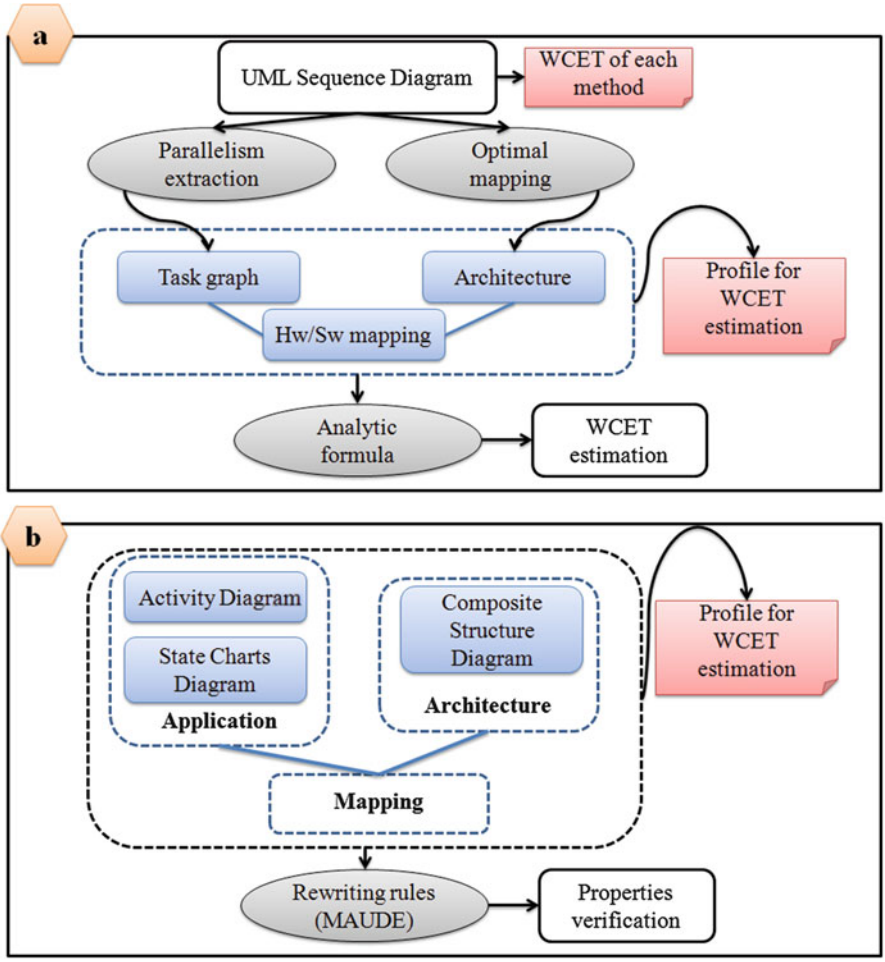

Fig. 3 Two UML-based design flows

These stereotypes allow annotating computations, communications, architecture components, and mapping.

Generating a task graph is guided by two steps: parallelism extraction from sequence diagram and optimal mapping. In fact, taking a sequence model as input, the designer can identify hierarchic, concurred, exclusive, pipelined, sequenced, and finally data partition behaviors. A manual extraction of the parallelism is then performed and a task graph can be drawn containing tasks and channels. These components are annotated with the provided stereotypes depending on the task type (behavior, sequence, etc.) and the communication type (signal passing and shared memory). The second step, the mapping, is based on some rules that deal with two constraints: decreasing communication overhead and minimizing the behaviors

execution time. These rules are used to identify the appropriate architecture for the application. The architecture is also specified manually in the same task graph and the mapping links are modeled between the computations and the hardware components. Now, having the application, the architecture, and the mapping, the WCET of the system can be estimated, thanks to some analytic formula.

The same author presents in [9] another UML-based approach for the modeling and formal verification for embedded systems (Fig. 3b). The application is considered a network of hierarchic data and control-based tasks communicating throughout abstract channels. While dataflow-based tasks are specified using the UML activity diagram, control-driven tasks are modeled with the UML State Charts diagram. Added to that, the structure diagram is used to specify the hardware components and the global structure of the architecture and different constraints are specified using UML constraints. These diagrams are enriched with stereotypes defined to guide the verification process. For the formal verification, the modeled diagrams including tasks, hardware components, and mapping are specified as sequence of rewriting rule using the MAUDE formal language. Using the MAUDE commands, the designer can verify some properties including deadlock, bus occupation rate, data size in FIFOs, etc.

There are currently a variety of profiles facilitating the modeling of embedded systems. In fact, several attempts have been proposed to adapt UML to real-time field such as UML-RT [52] and ACCORD/UML [31]. Meanwhile, the Object Management Group (OMG) worked on the standardization of two profiles that are SysML [38] and UML SPT [39]. With the ever increasing demand and complexity of embedded systems, a new profile has been emerged as an evolution of the UML SPT profile. This standardized profile named Modeling and Analysis of Real-Time and Embedded Systems (MARTE) [40] takes few concepts from the SysML modeling language.

Using MARTE stereotypes to specify two complex NoC topologies was the focus of the author in [15]. A powerful package of MARTE dedicated to model repetitive structures and topologies was examined to specify the Honeycomb [21] and the GEXspidergon [57] topologies. Generic models of the two NoCs are defined using the Repetitive Structure Modeling (RSM) package of MARTE and the routing algorithm is also modeled as a hardware component where the behavior is described with a UML state machine [16].

Parallel massive data processing is a key feature in the actual embedded systems. When dealing with massive computation and data-intensive processing, the use of massively parallel architectures is very useful. In 2010, a parametric massively parallel Single Instruction Multiple Data (SIMD) architecture, named mppSoC was proposed [34]. An mppSoC (massively parallel processing SoC) system is a generic massively parallel embedded architecture designed for data-parallel applications. mppSoC is designed based on an assembly of different components and may be implemented on a single chip. In addition, mppSoC proves very fruitful in massively parallel applications domain. However, the design and implementation of such systems become critical due to their long design and development cycles. In fact, the mppSoC design is facing today a strong pressure on reducing

time-to-market while the complexity of this system has been increasing. Changing one SoC configuration may also necessitate extensive redesign. Design abstraction offers a possible solution to address the above issues concerning the time-to-market and complexity dilemma. It is in the context of improving the primary productivity of mppSoC that the work of our team finds its proper place.

A first design approach for mppSoC was proposed in [4] (Fig. 4a). The presented mppSoC design flow uses the MARTE standard profile for high-level system specification. This flow is based on a Model-Driven Engineering (MDE) approach. It promotes separation of concerns, reusability, and automatic model refinement from higher abstraction levels to executable VHDL description.

The approach allows the designer to automatically select an mppSoC configuration at a very early design stage, before system synthesis and code generation have been performed. Our modeling methodology leverages from the MARTE profile the Hardware Resource Modeling (HRM), the RSM, and the Generic Component Model (GCM) packages.

The designed SIMD architecture is configurable and parametric; so that the designer can choose different configurations depending on the application requirements. UML2 templates [41] support mechanisms to express such characteristic. They are used, in this case, to easily define mppSoC parameters.

To generate code, a transformation flow is implemented. We have taken advantage of the various tools offered by Acceleo [37] to generate the corresponding VHDL code. An IP library is also defined to guide the code generation. This library gathers hardware IPs of the mppSoC architecture. Such transformation is called model-to-text transformation. As transformation chain's output, the user expects executable code, which can be used in already available tools. The target is a synthesize VHDL description for the mppSoC architecture.

A second tentative to facilitate the mppSoC design is described in [5] (Fig. 4b). In this approach, we tried to integrate our flow for code generation in the GASPARD [18] framework. The architecture is then specified using MARTE and the Deployment profile of GASPARD. The latter permits to select IPs for the modeled components from the IP library provided with the framework and enriched with IPs of mppSoC system. The code generation process is implemented as a refinement flow that takes high-level models as design entry and executes different transformation to generate VHDL code in the final step. The application can be described in C then compiled and introduced to the IP library. The binary file can be then deployed on the memory as a software IP in the deployment diagram of the specification.

In another work [2, 3], the mppSoC tool is proposed as an Eclipse plugin (Fig. 4c). This tool allows users to specify an mppSoC system, to explore the design space based on high-level analysis, and finally to automatically generate both the application and architecture source code. The entry point corresponding to the high-level specifications is a MARTE-compliant model. Such a model is specified by a user with UML modeling tools such as Papyrus [46]. A major goal of our tool is to rapidly design an mppSoC system that meets the needed requirements, in particular those related to performances. This goal is achieved by considering high-level

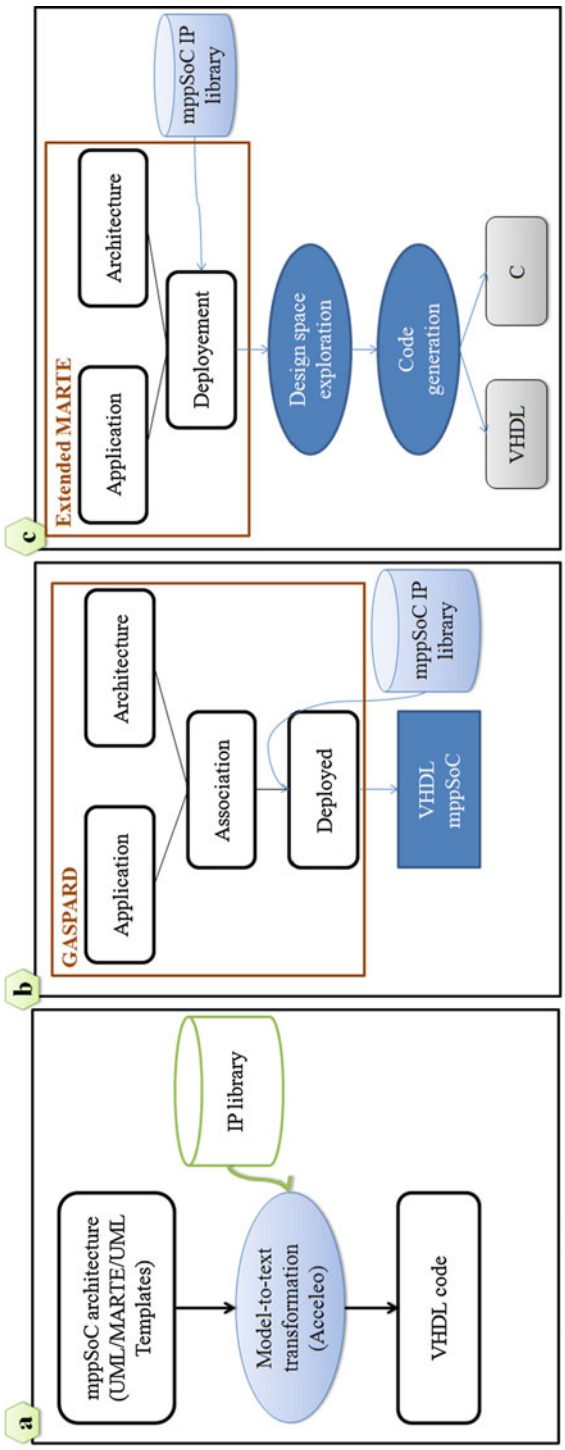

Fig. 4 The mppSoC design flows

models and the different configuration choices that are reached via the refinement chains provided in our framework. For that purpose, some precise ordered design steps should be respected. Based on these steps, we define a methodology dedicated to the high-level DSE for high performance mppSoC. The designer should exploit the stereotypes provided by the extended version of MARTE. So, he specifies the potential parallelism available in the different system parts: data-parallel application representing the software part and massively parallel architecture characterizing the hardware part. We distinguish four sub-steps that are combined in different ways:

- Description of system functionality and architecture: Here the user defines the application algorithm on the one hand, and the hardware architecture on the other hand. This step depends on the designer choices. If he wants to go through the DSE step, he must specify the data-parallel application and model the necessary parts of the architecture. But if he just wants to generate the application and architecture code without exploring design alternatives, he has to specify both the architecture by fixing all parameters related to his configuration. Our tool includes two extensions to model a complete mppSoC system. These extensions allow to model the mppSoC configuration and the data-parallel application.
- IP Deployment: At this stage, the elementary components used in the system functionality are deployed on IPs. If the designer choose to fix the mppSoC architecture parameters, the deployment of the architecture elementary components becomes necessary.
- DSE: The DSE step relies on two model transformations. These two transformations allow the exploration of a large design space via a multi-objective algorithm. Reaching a satisfactory solution refers to the selection of the appropriate solution. This selection is driven by the constraints fixed by the designer using the user interface.
- Application and architecture code generation: The final step in our co-design flow is the generation of the VHDL code of the mppSoC configuration and the C code of the data-parallel application. For this reason, two chains are defined. The first is the VHDL chain, which includes a model-to-text transformation. And the second is the C chain, which is based on two transformation types: a model-to-model and a model-to-text transformation.

5 Related Works

Recently, worldwide design methodologies have been suggested as an effective approach for embedded systems co-design. While some of these approaches support the DSE, other methodologies are mainly concentrated on the automatic code generation for the implementation.

The GASPARD [18] framework suggests an MDA-based approach for SoC co-design that mainly relies on meta-modeling concepts. These meta-models are used to model architecture, application, and software/hardware associations. It also

exploits two types of transformations (model-to-model and model-to-text) to obtain the executable implementation. A re-factoring mechanism is followed in order to find a better adequacy between model of the application and model of the architecture. The main focus of the GASPARD framework is the code generation for simulation at different abstraction levels. Among these levels we can mention the TLM and RTL levels.

The MODES framework presented in [14] uses the UML class, composite structure and sequence diagrams enriched with the MARTE stereotypes for the functional specification of the embedded application. It then automatically generates a control and data flow internal representation conforming to a number of meta-models. Additional meta-models and transformations are defined to perform the DSE phase. MODES is based on a flow of transformations between different tools. In fact, while the H-SPEX [43] and the SPEU [44] tools are used for the DSE and estimation tasks, the UPPAAL model checking tool is used to validate the functional and temporal properties of the specification, taking as input a network of timed automata.

The Sesame framework [12] offers capabilities for the modeling and simulation of heterogeneous embedded multimedia systems. While the functional behavior of the application is described using KPN, the architecture resources and their performance constraints are defined in the Pearl simulation language or in SystemC. The mapping layer is automatically generated, then a trace-driven simulation is executed to evaluate the system performances. The Artemis workbench [48] uses the Sesame framework as a basis for the performance analysis. The KPN application specifications are automatically generated by the Compaan tool-set starting from a Matlab-based description. Then, the Laura tool-set produces VHDL code taking as an entry the KPN models. The generated code implements the application for a specific FPGA platform. The Artemis framework uses specific platform architecture, the Molen platform, for the component calibration.

In the Koski flow [26], two steps are needed to ensure architecture exploration. First, a static analysis of the application model is done. Then iterative simulations take place in order to explore the architecture. This flow relies mainly on UML to model application, architecture, and the association between them.

COMPLEX [19] and MADES [49] are EU FP7 projects aiming to use high-level models for the design of real-time and embedded systems. The COMPLEX framework mixes system-level power optimization and rapid prototyping techniques in a PBD flow. Subsets of the MARTE profile are used to specify the system and UML use-cases are modeled to generate systems stimuli. While executable specifications are generated automatically from UML models allowing separate hardware and software estimation, the mapping is done manually and provided as a design entry. Several system-level estimation tools are combined together to facilitate power and execution time evaluation.

The MADES project proposes a methodology that combines MARTE and SysML in a design flow targeting embedded avionics systems. This flow focuses on modeling, verification, and hardware and software generation steps. However, the DSE step is not supported in this framework. In fact, the main focus of this

framework is the specification and how the user can combine different profiles to give a complete view of the system.

6 Discussion

6.1 Evaluation of African Research Efforts in the Co-design of SoCs

In this subsection, we attempt to compare the previously described African research efforts, to highlight the differences and similarities in their research directions based on the following table. The next subsection will compare the African research efforts and the worldwide approaches described in the related works section (Table 2).

First described approaches, language-based design flows, use the VHDL language to describe the architecture. From a behavioral specification, traditional design steps are executed (scheduling, allocation, hardware/software generation) at different levels of abstraction (behavioral or RTL). Moving from one level of abstraction to another is guaranteed by synthesis. While the software estimation is carried out using profiling or synthesis techniques, performances of the architecture are concluded after simulation. The performance estimation or the DSE phases are usually done at the RTL level using simulation or estimation metrics. The profiling technique is generally mixed with the partitioning phase into an iterative process to find the optimal design.

Table 2 Comparison of the described approaches

Approach	Base model	Focus	Target system
[1]	VHDL/C	DSE, synthesis, co-verification, prototyping	Simple model
[17]	C	Profiling, partitioning, performance analysis, co-verification, co-synthesis	Simple model
[6]	VHDL	Scheduling, allocation, architecture generation	ASIP
[51, 54, 30, 55]	TPG	Scheduling, partitioning, system generation	MPSoC
[53]	KPN/SDF	Performance estimation	MPSoC
[58]	KPN	Performance estimation using Sesame	MPSoC
[10]	UML	Manual mapping, performance estimation	MPSoC
[9]	UML/MAUDE	Formal verification	MPSoC
[4]	UML/MARTE	Architecture generation	mppSoC
[5]	UML/MARTE	Architecture generation with GASPARD	mppSoC
[2, 3]	UML/MARTE	DSE, architecture, and application generation	mppSoC

These flows provide either a complete system ready for the final integration or files of the hardware and software parts at a low-level of abstraction. The described flows target simple models of the architecture including one processing unit, simple memory architecture, and bus-based interconnections. Some of them are specifically created to generate code for ASIP or ASIC components.

Now, considering the age of model-based co-design flows, the major evolution that can be directly noticed is the raise in the levels of abstraction. System specification at that age is driven by graphs and UML diagrams. While graph-based flows give semantics to specify complex and high-performance applications, UML diagrams provide means to model the application, the architecture, and the mapping in one model. These high-level models facilitate the DSE and the hardware and software estimation tasks. Low-level simulation-based estimation techniques are no more employed. In fact, analytic formula, formal languages, and exploration algorithms are now exploited to facilitate the NP-complete exploration task, to reduce the time-to-market and to make the design costs lower. Partitioning and scheduling design phases become also easier at the model-level of abstraction.

Apart from these advantages, reducing the design complexity encourages the design of multiprocessor and even massively parallel SoCs. The design flows now are dedicated to complex architectures and most of them are based on the PBD approach making the design of complex architectures easier.

6.2 Recommendations to Improve the Competitiveness of Africa in SoC Design

Literature on worldwide co-design flows was reviewed in the previous section. The majority of these flows are the fruits of powerful projects like COMPLEX and MADES which are EU FP7 projects. Others are consistent co-design frameworks resulting from a long and hard team work. Compared with the previously detailed African co-design flows, these related works seem to be more relevant as they gather powerful research centers and laboratories to work on the same framework. African research efforts, in contrast, are individual efforts accomplished during thesis and not integrated into projects or team work. Encouraging team work, funding significant research projects and targeting bigger scale issues can improve the competitiveness of African researchers in SoC design.

6.3 Knowledge Transfer from Academia to Industry

The absence of communication between academia and industry makes R&D an insignificant word in the African research field and the African industry terminologies. However, the strong presence of R&D centers in the semiconductor and EDA

big companies shows that industry pioneers know deeply the benefit of academic research in the improvement of their product. As a result, the SoC design success relies on academic efforts and industry potentials. So, it is very important to create real research projects supervised by both academia and industry, trying to solve real problems. In this context, in Tunisia for example, a project to support research and innovation system (PASRI) was launched in 2012. The main goal is to ensure an effective networking and exchange between industry and universities in order to contribute in the development of Tunisia.

7 Conclusion

In this chapter, we survey the worldwide semiconductor and EDA industries evolution with a particular focus on the African industry. Our study demonstrates that the manufacturing leaders of the semiconductor industry are concentrated in Europe, USA, Taiwan, South Korea, and China. Some of these advanced countries have created few design centers in African countries. This minor transfer of core technologies to the African region aims to save production costs, while enhancing the design capabilities of the semiconductor and EDA industries. How to improve African industrial competitiveness and how to achieve good rank in the list of semiconductor and EDA companies are worthy questions to consider. We suggest that African countries take advantage of the opportunity to integrate companies from developed countries, while encouraging local firms to invest in semiconductor manufacturing and electronic tools design. Enhancing the R&D capabilities is another important solution since African companies in general do not rely on R&D teams. This will lead to improvements in the African industry by moving toward production instead of being consumers. Academic research efforts in the SoC design field were under investigation in a second part of this chapter. These efforts present several design flows that facilitate the design of systems to be implemented on integrated circuits. Our survey is to show their diversities in terms of focus, target system, and base model. For this reason, these flows are classified in two types: language-based approaches and model-based approaches taking the base model as classification criterion. The last part of our chapter is dedicated to evaluate these flows to come up with conclusions. Since massive research and industry efforts are being made to develop new embedded products and new design methods and tools for SoC design in the world, our eyes should stay wide open to ameliorate our position as African developers and academic researchers in the world.

References

1. M. Abid, Contribution à la conception des systèmes électroniques mixtes logiciel/matériel, Doctorat d'état, ENIM, 2000
2. M. Ammar, M. Baklouti, M. Abid, A model driven based CAD tool for mppSoC design, in *23rd International Conference on Microelectronics, ICM' 2011*, Hammamet, Tunisia, Dec 2011
3. M. Ammar, M. Baklouti, M. Abid, Extending MARTE to support the specification and the generation of data intensive applications for massively parallel SoC, in *15th Euromicro Conference on Digital System Design, DSD 2012*, 5–8 Sept 2012, pp. 715–722
4. M. Ammar, M. Baklouti, Ph. Marquet, M. Abid, J.-L. Dekeyser, A model driven engineering design approach to generate VHDL for MPPSoC, in *3rd Workshop on Rapid Simulation and Performance Evaluation: Methods and Tools, RAPIDO'2011*, (Heraklion, Crete, Greece, Jan 2011)
5. M. Baklouti, M. Ammar, P. Marquet, M. Abid, J.-L Dekeyser, A model-driven based framework for rapid parallel SoC FPGA prototyping, in *22nd IEEE International Symposium on Rapid System Prototyping, RSP 2011*, May 2011, pp. 149–155
6. M. Benmohammed, S. Merniz, M. Bourahla, ASIP micro-code generation from high-level specifications, in *Proceedings of the International Conference on Information and Communication Technologies: From Theory to Applications*, 19–23 April 2004, pp. 587–588
7. M. Benmohammed, A. Rahmoune, P. Kission, The application of high-level synthesis techniques for the generation of pipelined reprogrammable microcontrollers, in *The 7th IEEE International Conference on Electronics, Circuits and Systems, ICECS 2000*, vol 2, (2000), pp. 993–998
8. F. Boutekkouk, M. Benmohammed, S. Bilavarn, M. Auguin, UML for modelling and performance estimation of embedded systems. J. Object Technol. **8**(2), 95–118 (2009)
9. F. Boutekkouk, M. Benmohammed, UML modeling and formal verification of control/data driven embedded systems, in *14th IEEE International Conference on Engineering of Complex Computer Systems*, 2–4 June 2009, pp. 311–316
10. F. Boutekkouk, S. Bilavarn, M. Auguin, M. Benmohammed, UML profile for estimating application worst case execution time on system-on-chip, in *International Symposium on System-on-Chip, SOC 2008*, 5–6 Nov 2008, pp. 1–6
11. A. Charfi, A. Gamatié, A. Honoré, J.L. Dekeyser, M. Abid, Validation de modèles dans un cadre d'IDM dédié à la conception de systèmes sur puce," 4ème Journées sur l'Ingénierie Dirigée par les Modèles (IDM 08) (2008)
12. J.E. Coffland, A.D. Pimentel, A software framework for efficient system-level performance evaluation of embedded systems, in *Proceedings of the ACM Symposium on Applied Computing (SAC)*, Mar 2003, pp. 666–671
13. J.E. Coffland, A.D. Pimentel, A software framework for efficient system-level performance evaluation of embedded systems, in *Proceedings of the ACM Symposium on Applied Computing*, Mar 2003, pp. 666–671
14. F.A.M. do Nascimento, M.F.S. Oliveira, F.R. Wagner, A model-driven engineering framework for embedded systems design. Innov. Syst. Softw. Eng. **8**, 19–33 (2012)
15. M. Elhaji, P. Boulet, S. Meftali, A. Zitouni, J.-L. Dekeyser, R. Tourki, An MDE approach for modeling network on chip topologies, in *5th International Conference on Design and Technology of Integrated Systems in Nanoscale Era, DTIS 2010*, 23–25 Mar 2010, pp. 1–6
16. M. Elhaji, P. Boulet, R. Tourki, A. Zitouni, J.-L. Dekeyser, S. Meftali, Modeling networks-on-chip at system level with the MARTE UML profile, in *M-BED 2011* (Grenoble, Mars 2011)
17. M.W. El-Kharashi, M.H. El-Malaki, S. Hammad, A. Salem, A. Wahdan, Towards automating hardware/software co-design, in *Proceedings of the 4th IEEE International Workshop on System-on-Chip for Real-Time Applications*, 19–21 July 2004, pp. 189–192
18. A. Gamatié, S. Le Beux, E. Piel, R. Ben Atitallah, A. Etien, P. Marquet, J.L. Dekeyser, A model-driven design framework for massively parallel embedded systems. ACM Trans. Embed. Comput. Syst. (TECS) **10**, 39 (2011)

19. K. Grüttner, P.A. Hartmann, K. Hylla, S. Rosinger, W. Nebel, F. Herrera, E. Villar, C. Brandolese, W. Fornaciari, G. Palermo, C. Ykman-Couvreur, D. Quaglia, F. Ferrero, R. Valencia, The COMPLEX reference framework for HW/SW co-design and power management supporting platform-based design-space exploration. Microprocess. Microsy. **37**, 966–980 (2013)
20. F. Hannig, H. Dutta, A. Kupriyanov, J. Teich, R. Schaer, S. Siegel, R. Merker, R. Keryell, B. Pottier, O. Sentieys, Co-design of massively parallel embedded processor architectures, in *Proceedings of the 1st Reconfigurable Communication-Centric Systems on Chip workshop, ReCoSoC'05*, Montpellier, France, June 2005
21. A. Hemani, A. Jantsch, S. Kumar, A. Postula, M. Millberg, D. Lindqvist, Network on chip: an architecture for billion transistor era, in *Proceedings of the IEEE NorChip Conference*, Nov 2000
22. R. Hocine, H. Kalla, S. Kalla, C. Arar, A methodology for verification of embedded systems based on SystemC, in *IEEE International Conference on Complex Systems (ICCS)*, Nov 2012, pp. 1–6
23. International Technology Roadmap for Semiconductors, Modeling_2012 Tables
24. ITRS Design Technical report, International technology roadmap for semiconductors, (2011)
25. G. Kahn, The Semantics of a Simple Language for Parallel Programming, IFIP Congress, 1974, pp. 471–475
26. T. Kangas, P. Kukkala, H. Orsila, E. Salminen, M. Hännikäinen, T. Hämäläinen, UML-based multiprocessor SoC design framework. ACM Trans. Embe. Comput. Syst. **5**, 281–320 (2006)
27. K. Keutzer, A.R. Newton, J.M. Rabaey, A. Sangiovanni-Vincentelli, System-level design: orthogonalization of concerns and platform-based design. IEEE Trans. Comput-Aided Des. Integr. Circuits Syst. **19**(12), 1523–1543 (2000)
28. P. Kission, H. Ding, A.A. Jerraya, Structured design methodology for high-level design, in *Proceedings of the 31st ACM/IEEE DAC*, June 1994
29. H. Krichene Zrida, A. Jemai, A.C. Ammari, M. Abid, High level H.264/AVC video encoder parallelization for multiprocessor implementation, in *12th ACM/IEEE Design Automation and Test in Europe conference and exhibition, DATE '09*, 20–24 Apr 2009, pp. 940–945
30. Y.-K. Kwok, I. Ahmad, Static scheduling algorithms for allocating directed task graphs to multiprocessors. ACM Comput. Surveys (CSUR) J. **31**(4), 406–471 (1999)
31. P. Lanusse, S. Gérard, F. Terrier, Real-Time Modeling with UML: The Accord Approach, The Unified Modeling Language. «UML»'98: Beyond the Notation, vol. 1618, (1999), pp. 319–335
32. E. Lee, D. Messerschmitt, Synchronous data flow. Proc. IEEE **75**(9), 1235–1245 (1987)
33. G. Liu, *More-than-Moore Reinvented, International Symposium on Semiconductor Manufacturing (ISSM)*, Oct 2010, pp. 1–18
34. M. Baklouti, A rapid design method of a massively parallel System on Chip: from modeling to FPGA implementation, http://tel.archives-ouvertes.fr/tel-00527894/en/
35. Mentor Graphics, Application Specific Assistant Processor (ASAP), 2003
36. G.E. Moore, Cramming more components onto integrated circuits. Proc. IEEE **86**(1), 82–85 (1998)
37. Obeo. Acceleo Generator, http://www.acceleo.org (2014)
38. Object Management Group, UML Profile for Schedulability, Performance, and Time, version 1.1. http://www.omg.org/spec/SPTP/1.1
39. Object Management Group, The OMG Systems Modeling Language (SysML), Version 1.2. http://www.omg.org/spec/SysML/1.2/PDF
40. Object Management Group, UML Profile for MARTE: Modeling and Analysis of Real-Time Embedded Systems, version 1.0. http://www.omg.org/spec/MARTE/1.0/PDF/
41. Object Management Group, Inc. (2005) UML Technical Report formal/2005-07-04
42. Object Management Group. Unified Modeling Language specification, version 2.1. http://www.omg.org/spec/UML/2.1.2/Infrastructure/PDF
43. M.F.D.S. Oliveira, E.W. Brião, F.A.M.D. Nascimento, F.R. Wagner, Model driven engineering for MPSOC design space exploration, in *Proc. SBCCI*, 2007, pp. 81–86
44. M.F.S. Oliveira, L.B. de Brisolara, L. Cairo, F.R. Wagner, Early embedded software design space exploration using UML-based estimations, in *the 17th IEEE International Workshop on Rapid System Prototyping, RSP 2006*, Chania, Crete, June 2006

45. OMG, Object Constraint Language OMG Available Specification, Version 2.0 formal/06-05-01
46. Papyrus, 2014, http://www.eclipse.org/papyrus/
47. H.-w. Park, H. Oh, S. Ha, Multiprocessor SoC design methods and tools. IEEE Signal Proc. Mag. **26**(6), 72–79 (2009)
48. A.D. Pimentel, The Artemis workbench for system-level performance evaluation of embedded systems. Int. J. Embe. Syst. **3**(3), 181–196 (2008)
49. I.-R. Quadri, E. Brosse, I. Gray, N. Matragkas, L.S. Indrusiak, M. Rossi, A. Bagnato, A. Sadovykh, MADES FP7 EU project: Effective high level SysML/MARTE methodology for real-time and embedded avionics systems, in *7th International Workshop on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC)*, July 2012, pp. 1–11
50. IHS iSuppli Research, Semiconductor industry revenue to ensure slow growth in 2012. (2014), http://www.isuppli.com/Semiconductor-Value-Chain/News/Pages/Semiconductor-Industry-Revenue-to-Endure-Slow-Growth-in-2012.aspx
51. M. Safar, M.A. El-Moursy, M. Abdelsalam, A. Salem, Architecture exploration of multicore systems-on-chip using a TLM-based framework. IJCSI Int. J. Comput. Sci. Issues **10**(4), 2 (2013)
52. B. Selic, Using UML for modeling complex real-time systems, in *Proceedings of the ACM SIGPLAN Workshop on Languages, Compilers, and Tools for Embedded Systems, LCTES '98*, London, 1998, pp. 250–260
53. K. Smiri, A. Jemai, NoC–MPSoC performance estimation with synchronous data flow (SDF) graphs, Lecture Notes in Computer Science, (Springer, 2011), pp. 406–415
54. K. Yehia, M. Safar, H. Youness, M. AbdElSalam, A. Salem, A design methodology for system level synthesis of multi-core system architectures, in *Saudi International Electronics, Communications and Photonics Conference (SIECPC)*, 24–26 Apr 2011, pp. 1–6
55. H. Youness, M. Hassan, K. Sakanushi, Y. Takeuchi, M. Imai, A. Salem, A.-M. Wahdan and M. Moness, A high performance algorithm for scheduling and hardware–software partitioning on MPSoCs, in *4th International Conference on Design & Technology of Integrated Systems in Nanoscal Era, DTIS '09*, 6–9 Apr 2009, pp. 71–76
56. H. Youness, M. Hassan and A. Salem, A Design Space Exploration methodology for allocating Task Precedence graphs to multi-core system architectures, in *International Conference on Microelectronics, ICM 2010*, 19–22 Dec 2010, pp. 260–263
57. M. Zid, A. Zitouni, A. Baganne, R. Tourki, Nouvelles architectures génériques de NoC = news generics architectures for NoC. Technique et Science Informatiques **28**(1), 101–133 (2009)
58. H.K. Zrida, A.C. Ammari, A. Jemai, M. Abid, System-level performance evaluation of a H.264/AVC encoder targeting multiprocessors architectures, in *International Conference on Microelectronics, ICM 2009*, 19–22 Dec 2009, pp. 169–172

New Coarse-Grained Configurable Architecture for DSP Applications

Heni Ghariani, Zied Marrakchi, Abdulfattah Obeid, Mohammed S. BenSaleh, Mohamed Abid, and Habib Mehrez

Abstract While observing the evolution of reconfigurable architectures, one can see how the designers' experience and the market demand factors reshaped the newest reconfigurable FPGAs products. Fine-grained FPGA architectures provide high flexibility at the cost of area, performance, and power degradation. There is a trend to use coarse-grained architecture for some specific applications requiring high computational resources. In this chapter we present a study of a coarse-grained reconfigurable architecture capable of effectively implementing specific DSP applications. We aim to achieve a good tradeoff between flexibility and performance/density. Relying on the study of different DSP applications and their corresponding VLSI architectures, we propose a generic reconfigurable architecture capable of realizing the studied DSP functions. The logic block is composed essentially of an adder and a multiplier. The logical units (elements of arithmetic treatment and elements memorizing) are arranged in lines and columns forming a generic matrix and interconnected by a configurable network (direct connections and/or programmable bus). CAD tools are proposed to automate the implementation of the design on the architecture. To validate and further exploit the aforementioned architecture, a synthesizable and parametrizable VHDL model generator was developed. This VHDL model generator allowed experimenting with different architectures facilitating the optimization of the efficiency of the implemented designs. Thanks to unifying coarse-grained logic block and bus-based interconnections we reduce the gap between ASIC and configurable architecture area from 38 to 7 times.

H. Ghariani • M. Abid (✉)
CES Laboratory, National Engineering School of Sfax, Sfax, Tunisia
e-mail: heni.ghariani@yahoo.fr; mohamed.abid_ces@yahoo.fr; mohamed.abid@enis.rnu.tn

Z. Marrakchi • H. Mehrez
Laboratoire LIP6-SoC, University of Paris VI, Pierre & Marie Curie, Paris, France
e-mail: Zied.Marrakchi@lip6.fr; Habib.Mehrez@lip6.fr

A. Obeid • M.S. BenSaleh
The National for Electronics, Communications and Photonics Center, King Abdulaziz City for Science and Technology, Riyad, Kingdom of Saudi Arabia
e-mail: obeid@kacst.edu.sa; mbensaleh@kacst.edu.sa

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_10

1 Introduction

Several African countries carry out research in microelectronics and their applications on modern products with an emphasis on high technologies. We strive for academic excellence and research relevance on this domain through various research, development, and innovation. Tunisia in collaboration with other countries has implemented a range of innovative surveys projects in universities. Most of these projects facilitated African researchers to master and transfer high-level technologies. In this context, the work presented in this chapter is carried out and is performed in cooperation between:

- CES Lab: Computer Embedded System, National School of Engineering of Sfax, Tunisia
- LIP6: Laboratory of Computer Science, Paris 6, France
- KACST: King Abdulaziz City for Science and Technology, Riyadh, KSA

The goal is to propose a new coarse-grained reconfigurable architecture capable to implement effectively specific DSP applications. In fact, between the two computational paradigm extremes with the microprocessor trading computational efficiency for flexibility and ASICs trading area and performance efficiency there is clearly a design space gap that can be occupied by Reconfigurable Computing (RC) [6, 8, 29]. By dedicating more general purpose hardware resources that can be interconnected and configured to realize different functions, RC solutions aim at finding balance between pure software and pure hardware approaches. This is done through hardware structures that can be reconfigured to carryout different tasks at different times. For all its flexibility and performance, RC presents a promising solution not only for closing the various design gaps, but moreover, for bringing about a multitude of computational possibilities. Due to several technological challenges, the full capacity of RC is yet to be achieved. On the other hand, there are self-induced design challenges resulting from the very nature of reconfigurable computing: reconfigurability & computation. From a computational point of view having a very powerful array of processors capable of very high throughputs may not be useful at all if there is no feasible way to use it. Software support and integration along with other system issues present true bottlenecks. From the reconfiguration point of view new concepts such as Run-Time-Reconfiguration operation, partial reconfigurability, and speed of reconfiguration are among the new concepts that have emerged as a byproduct of reconfigurable technologies. Finding efficient design schemes addressing the abovementioned concepts has triggered quite an amount of research.

Reconfigurable Architectures (RA) are generally composed of an array of Processing Elements (PE). These PEs can be as small as a 4-input 1-output Look-Up Table (LUT) or as big as a 16-bit Arithmetic Logic Unit (ALU). The term granularity refers to the size (and type) of the inputs and outputs of the processing elements. In the case of a PE consisting of a 4-input LUT the reconfigurable architecture is classified as a fine-grained architecture (FGA) and a RA array composed of 16-bit

ALU PEs can be classified as a coarse-grained architecture (CGA). FGA solution is capable of realizing a quite wide range of designs, but this capability comes at the expense of expensive routing and reconfiguration resources [25]. CGA solutions on the other hand are better suited for data flow designs but are not suitable for designs that involve control dominated or irregular fine-grained operations.

The challenges facing the development of Reconfigurable Computing fall mainly in two categories: classical and self-induced. Classical problems lie in the higher consumption of area and power as well as slower operation as compared to finely crafted Application Specific Integrated Circuits (ASICs). These problems are being solved in research mainly by finding optimal PEs that exhibit high functionality with smaller area and lower power consumption. In the following, we will introduce a CGA capable of efficiently realizing a family of DSP applications. We seek to reach good flexibility and reduction of development and integration costs. Building on observations of different DSP applications and their implemented VLSI architectures, we propose a CGA with parameterizable and flexible blocks. The generated CGA is designed to efficiently realize the DSP function of interest.

In the past few decades, developments in field of Digital Signal Processing (DSP) combined with the advancements in process technology and computer arithmetic techniques have enabled the realization of a magnitude of products and application that were not previously possible to realize. [4]. A straight forward approach is to implement the DSP algorithm of interest in software and have it run on a microprocessor or a Digital Signal Processor. Despite of their evolution, Microprocessors and Digital Signal Processors do not meet the requirements of many modern applications that are more complex and require more performance and low power consumption.

In this context, specialized ASIC solutions capable of delivering the required computations at the required constraints are utilized, thus yielding an increase in orders of magnitude of performance. This is a solution that lacks flexibility and involves increasing amounts of development costs. Additionally, large ASICs can take a year or more to design. Field Programmable Gate Arrays (FPGA) may be a solution. One great advantage of an FPGA is its flexibility and reconfigurability. CGA solutions are good candidates to be embedded in DSP applications.

In order to improve FPGA efficiency for DSP applications, a coarse-grained architecture approach can be adopted. The architecture features are suited to specific digital signal processing algorithms. This solution is composed of arithmetic and memory processing elements (PE). These PEs are arranged in rows and columns forming a 2D mesh. The PEs are interconnected by a configurable network. We define a tile-based architecture that can be duplicated by duplicating the same tile several times; thus saving development and verification time.

The remainder of this chapter is organized as follows: the next section presents some Coarse-Grained Reconfigurable Computing (CGRC) solutions available in the market or in the literature. This is followed by a general discussion on the previewed architectures and their different features and our observation of their

development trends. In Sect. 3 we discuss generic and flexible elements that are the basic building blocks of the proposed architecture. Section 4 summarizes realizations and validation of proposed technique. Some CAD tools are added in order to skill and accelerate the design on the one hand and mapping several operators and some DSP algorithm on the other hand. This leads to illustration examples of several DSP algorithms. Finally the chapter is concluded in Sect. 5.

2 Introduction to State of the Art CGA Solutions

Several CGA solutions are reported in the literature or available in the market. They are characterized according to their general organization, their Processing Elements (PE) architectures, their interconnect structures, and their reconfiguration characteristics. In the following some projects examples.

2.1 The Kress Array

The Kress Array is a dynamically reconfigurable regular array of 32-bit reconfigurable Data-Path Units (rDPUs) [10, 11]. The operation of rDPUs is data driven. The Kress Array went through several phases of development. The rDPU is the basic processing element of the Kress Array. Each processing element is capable of all the basic integer C language operations support where simple operations are carried out directly and more expensive ones such as multiplication and division are carried out in a micro programmed sequence style. rDPUs can also be used for routing data through them. An rDPU consists of an ALU, a register file, and a number of multiplexers facilitating full connectivity. The register file can be used to store constants, intermediate results, or frequently used inputs. The operations of the rDPU are data driven and are carried out independently of the rest of the array. There are three levels of interconnects in the Kress Array. In the bottom most level data is transferred between rDPUs to propagate and process intermediate results. Connection by abutment between rDPUs simplifies routing and is more suited to coarse-grained applications.

Global busses facilitate long distance connectivity between rDPUs and/or the higher level input/output busses. The input/output busses are interfaced with the internal global busses through switches. This style of hierarchical bus routing allows input data transferred to and from rDPUs not located at the edge of the array. Figure 1 illustrates a 9-rDPU Array structure. In order to reduce the number of input/output pins, serial mode connections are provided between local interconnects between sub-arrays. This serial connectivity is transparent to the programmer.

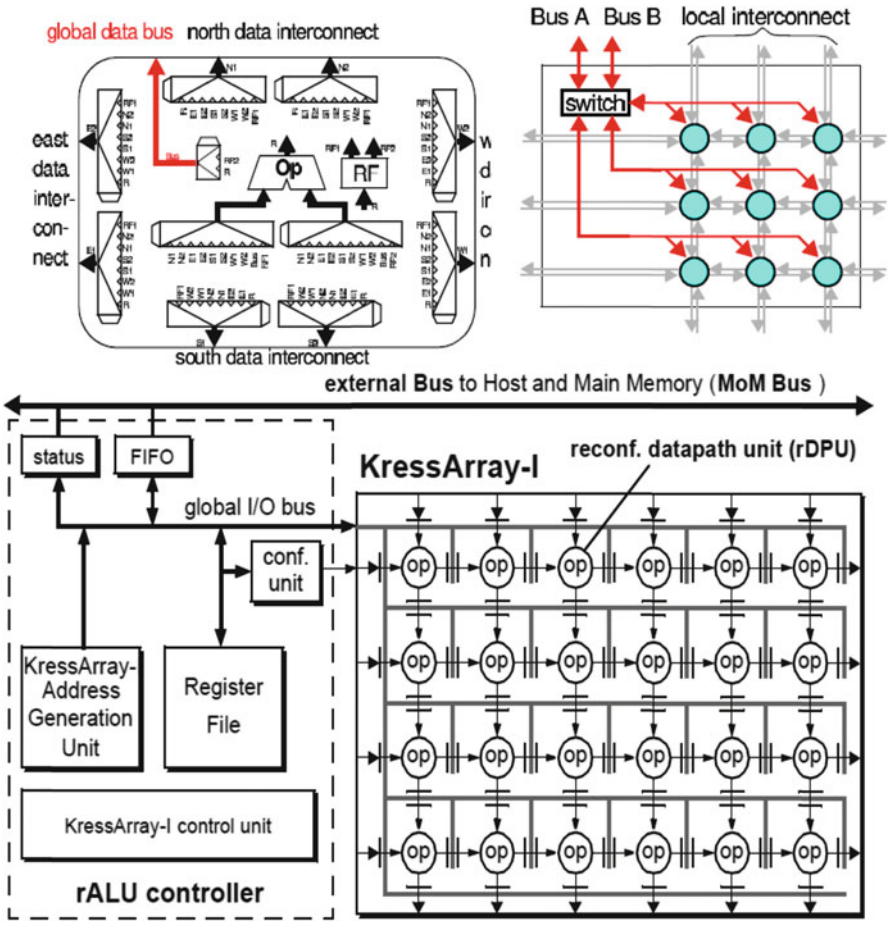

Fig. 1 Kress Aray architecture

2.2 The Reconfigurable Multimedia Array Coprocessor

The Reconfigurable Multimedia Array Coprocessor (REMARC) was introduced in [2]. The architecture is a reconfigurable coprocessor aimed at multimedia applications like video compression, video decompression, and image processing. The REMARC is tightly coupled to a MIPS- II ISA RISC processor, which supplies support for coprocessors. The processor has been extended by special instructions to access the REMARC architecture in the same way as a floating point coprocessor is accessed.

For example, a REMARC consists of an 8 by 8 array of processing elements (also called Nano processors), which is attached to a global control unit. The control unit manages data transfers between the main processor and the reconfigurable array

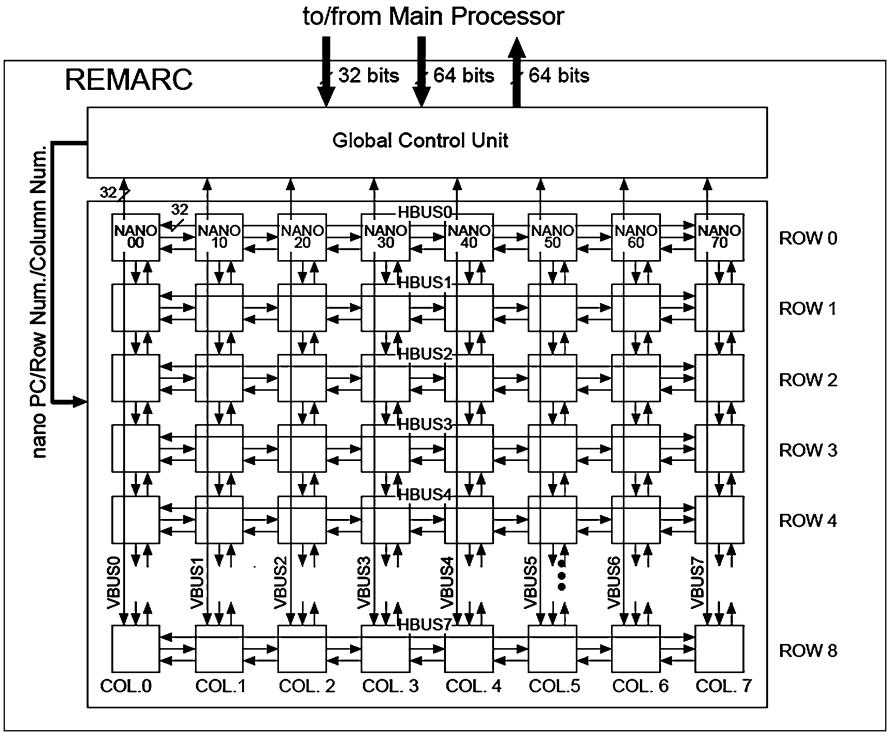

Fig. 2 Reconfigurable multimedia array coprocessor

and controls the execution of the Nano processors. It comprises an instruction RAM with 1,024 entries, 64 bit data registers, and four control registers.

The Nano processors consist of a local instruction RAM with 32 entries, an ALU with 16 bit data path, a data RAM with 16 entries, an instruction register, eight data registers, four data input registers, and one data output register. The ALUs can execute 30 instructions, including add, subtract, logical operations, shift instructions, as well as some operations often found in multimedia applications like minimum, maximum, average, and absolute. Each ALU can use data from the data output registers of the adjacent processors via nearest neighbor connect, from a data register, a data input register, or from immediate values as operands. The result of the ALU operation is stored in the data output register. The communication lines consist of nearest neighbor connections between adjacent Nano processors and additional horizontal and vertical buses in each row and column. The nearest neighbor connections allow the data in the data output register of a Nano processor to be sent to any of its four adjacent neighbors. The horizontal and vertical buses have double width (32 bits) and allow data from a data output register to be broadcast to processors in the same row or column respectively. Further, the buses can be used to transfer data between processors, which are not adjacent to each other. As can be seen in Fig. 2, the eight vertical buses are also used to connect the processor array to

the global control unit and to transfer data from or to the 64 bit data registers inside the controller. Hereby, the 64 bit data word can be split up onto the eight buses.

During execution, the global control unit issues a global program counter value each cycle, which is broadcast to all Nano processors. Each Nano processor then uses this global value to index its local instruction memory and execute the corresponding operation. To support SIMD operations for multimedia applications, it is possible to configure a whole row or column of processors to execute the same operation with a single instruction, thus saving instruction memory in the Nano processors.

2.3 The HPAD

The HPAD (Heterogeneous PAD) was introduced with its design influenced by the study of VLSI architectures of DSP functions [23]. The core of the HPAD is a Data Path Array (DPA) consisting of two types of PEs interleaved and interconnected in a "woven" torus topology. Around the DPA rest a number of peripherals to help manage operations and configurations. The DPA has mainly two types of PEs: an arithmetic and logic PE capable also of MAC operations and a Memory manipulation PE capable of local storage and simple bit and vector FIFO operations. Both units are interleaved and this interconnected.

Each PE is interconnected with a number of PEs around it and that include a number of butterfly connections which are very common in DSP applications. Additionally, long range vertical and horizontal busses spanning the complete DPA facilitate flexibility in routing data which results in efficient use of the resources on the HPAD. The selection of the types and functions of the PEs and their interconnectivity was inspired by the study of VLSI implementation of several DSP applications. Figure 3 depicts the basic organization of the DPA and its simplified routing topology.

To facilitate Run-Time-Reconfiguration (RTR) and partial reconfiguration, each PE is embedded in a socket carrying its current and possibly next context word. The context word contains information about the task and routing of this PE. The context can be switched at runtime depending on the function it performs, the result of a given result, a time stamp signal, or global change in the operation. A time keeper provides globally timing information needed to synchronize operations in some functions. Background reconfiguration is also allowed on the inactive contexts.

A configuration manager orchestrates the operation in general and configuration of the DPA. It also caters for signals coming from outside the HPAD informing it if data is ready to be supplied or consumed and also the status of the input and output FIFOs. Accordingly, the processing pipeline can be frozen of activated depending on the status of the FIFOs. The configuration manager also manages the configuration of the DPA either in the background or in the foreground.

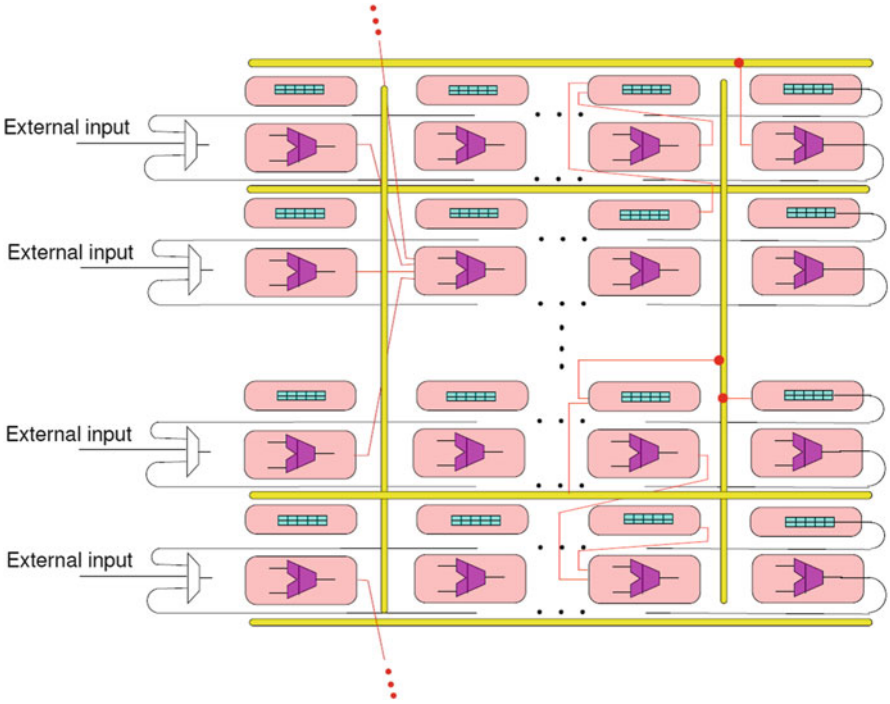

Fig. 3 A simplified illustration showing the routing topology of the HPad DPA

2.4 VPR Heterogeneous Architecture

VPR architecture (Fig. 4) and set of Tools are widely utilized in industrial and academic research projects. The first proposed architecture corresponds to an Island Style mesh architecture where LUTs are surrounded by configurable interconnect [5]. Interconnecting resources were based on bi-directional tri-states.

The architecture evolved from fine-grained LUT-based architecture [5] to heterogeneous architecture [18]. Today VPR architecture includes LUTs with heterogeneous set of logic blocks as adders and multipliers.

In addition, the configurable interconnect moved from tri-state-based bi-directional connections to unidirectional multiplexer-based connections. Nevertheless, the architecture does not operate in multi-bit buses of data. Thus, the fine-grained interconnect is very penalizing in terms of area for data flow applications style.

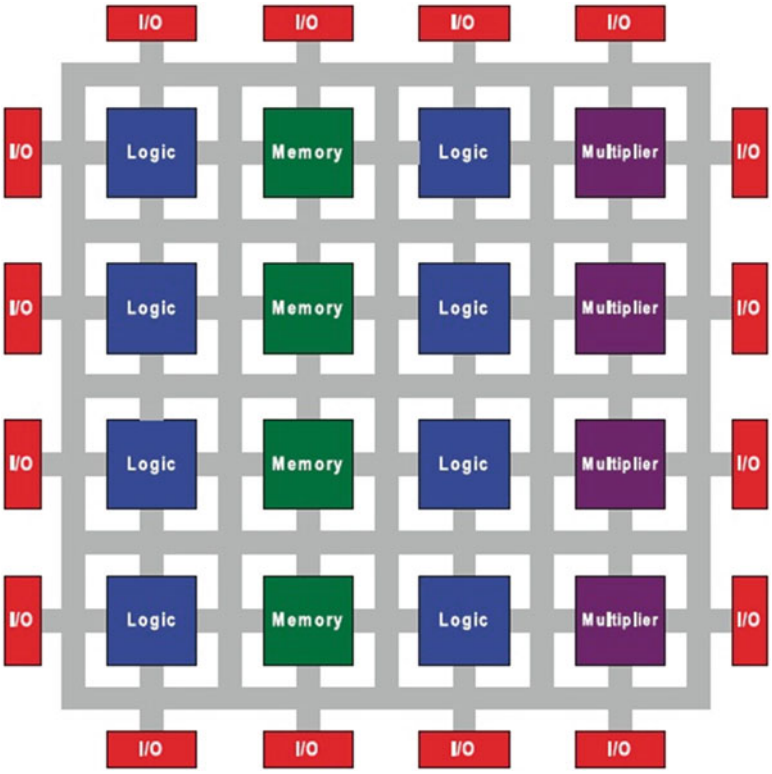

Fig. 4 VPR architecture

2.5 Concluding Remarks

Several other solutions and architectures of different flavors are available in the literature and in the market. They exhibit different routing, data flow, and configuration concepts. Amongst which are the MATRIX [22], MorphoSys [19], RAW [24, 27]. The reader is also referred to [1, 3, 28] for further examples.

Traditionally embedded computing tasks have been handled by ASICs as well as FPGAs. These provide two good alternatives to embedded processing in the areas of power, speed, and configurability. However, they are still only two regions in a large multi-dimensional space. FPGAs provide fine-grained parallelism with a degree of performance much better than processors and DSPs. ASICs have higher clock speeds and have lower power requirements, but lack the flexibility of an FPGA. Still, there are regions in the design space that can be explored that fit between ASICs and FPGAs. Architectures described in this section aimed to narrow the gap between FPGAs and ASICs by proposing a coarse-grained FPGA architecture (CGA) at both levels: Logic blocks or routing connectivity. We propose to develop an architecture

shifting away from LUT-based logic, which is inherently fine-grained, to logic units that operate on coarse-grained data and use bus-based connections to make routing. In this way we restrict our architecture to a set of signal processing applications to reduce cost and thus improve area, performance, and power consumption while maintaining good flexibility.

3 Flexible Coarse-Grained Architecture

Clearly, each computational paradigm is best suited for a given range of tasks. Irregular tasks and those with complicated control flow are best solved by microprocessor when the speed of operation is not critical. When very efficient realizations in terms of both performance and power consumption are intended, ASICs are the most appropriate especially when flexibility is not an important issue. For tasks that require good performance, power consumption and flexibility RC solutions present the best suited solutions.

In fact, the nature of DSP applications is extensive processing of data. Within a given algorithm as the studied Signal Flow Graphs suggest the number of computations explodes. Hardware implementation is investigated in area to increase the performance. The DSP algorithm can be implemented either as a parallel architecture or as pipelined architecture. In the parallel architecture approach parallelism in the DSP application is exploited and usually blocks of data are processed and produced at each clock cycle. This approach usually needs minimal control since data is consumed, processed, and produced in parallel with no need for intermediate storage and operation scheduling because the complete Signal Flow Graph (SFG) is implemented and most of the operation can be embedded in the structure and connectivity. However parallel architectures are very expensive in terms of area requirements and thus are only used when high performance is critical. In the so called pipelined approach data is consumed and produced serially. Therefore, scheduling of intermediate data and operations is required to allow operation on the appropriate pair (or set) of data supplied to different computational elements. The control resources needed are usually of moderate size especially when the SFG is regular. Area requirements of both data path and control resources are typically considerably lower than that of the parallel approach. In general, pipelined architectures are widely used and exhibit adequate performance. Thus, pure parallel implementations result in very high throughputs at the expense of enormous hardware resources especially for bigger sized algorithms. On the other hand, pipelined architectures intend to very efficiently trade speed for hardware resources of parallel architectures. This is done by unrolling and re-rolling their SFGs swapping therein intermediate data in dynamic data paths.

As our goal was to reach an efficient CGA solution for DSP applications and based on the aforementioned observations we decided that a pragmatic methodology to go about the design task is to study several DSP algorithms and their reported implementations extracting common and essential features to be considered when

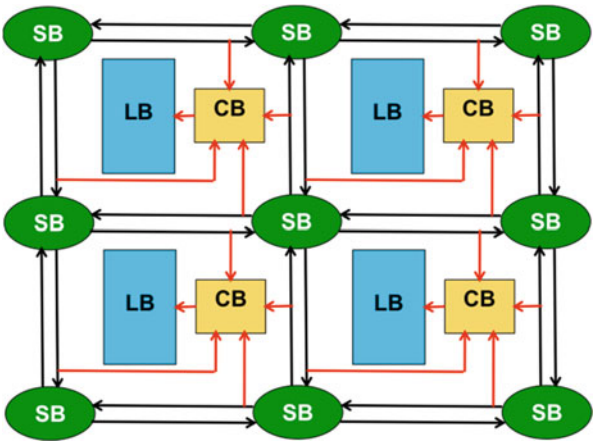

Fig. 5 2 × 2 matrix

designing our proposed CGA architecture. In fact, applications oriented data, when supporting different algorithms, uses typical signal processing operations. These common functions can be identified and then exploited to take advantage from the commonalities among common tasks in order to enhance power efficiency and area occupation. In this context, parameterization techniques can be used in order to identify the common aspects among different algorithms in order to define a generic operation capable of handling required tasks. This generic operation can switch from a configuration to another by a simple change of its parameters.

In fact, the proposed architecture is assimilated to a matrix which is composed of similar blocks. These blocks are called tiles. The tile is the smallest block, which is duplicated multiple times to build the required architecture. The advantage of having redundant tiles is the gain in development time. It is easier to design a small part of the architecture than replicated to build the overall architecture. In addition, if a modification is needed, it is sufficient to modify one block and then duplicating it. Also, this technique of duplication gives the user the possibility to define the size of the matrix. As a result more flexibility is given to the user. Three main blocks constitute the proposed CGA: The Logic Block (LB), The Connection Block (CB), and The Switch Block (SB). Figure 5 illustrates an example of Matrix.

3.1 Configurable Logic Block Architecture

The Logic Block acts as an Arithmetic unit for the proposed architecture. After studying the DSP algorithms, we deduce that most used operations are addition and multiplication [9]. Starting from this, the logic block is composed by an adder and a multiplier. The functions supplied by the logic block are: multiplication, addition, and addition followed by multiplication. So, to ensure data flowing between the two blocks some multiplexers are needed. Two function modes are carried out: the

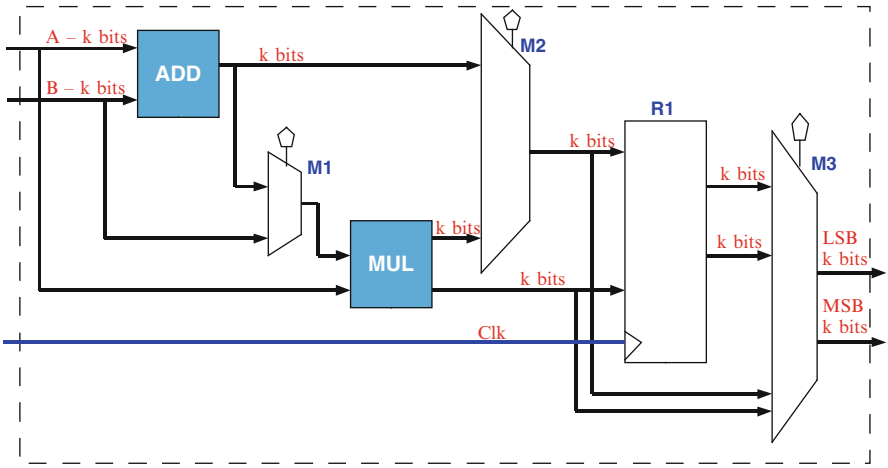

Fig. 6 The logic block architecture

sequential mode (with a register) and the combinatorial mode. The logic block takes two n-bits bus inputs and produces two outputs (LSB and MSB) of n-bits each one. As shown in Fig. 6, the elementary operations handled by the logic block are:

Adder: Each adder is composed of the combination of two half adders. The n-bits adder is composed by a succession of 1 bit adder. This architecture ensures the carry propagation. Each 1 bit adder has three inputs; the two operands and the carry bit are generated from the previous adder and generate two outputs: the result bit and the carry bit. In order to have an n-bit adder it is sufficient to cascade as desired 1 bit adder block. Each two bits of the same level from the inputs are added together and the carry provided from pervious adder block. For the Last Significant Bit the carry is a logic zero.

The Multiplier: The multiplier is composed of a succession of full adders. An n bit multiplier accepts two inputs and provides one output. The size of each input is n and the size of the output is 2n bits. For the proposed architecture the output is divided into two parts, the Less Significant Bits (LSB) and the Most Significant Bits (MSB). This division is needed since the output of the multiplier can be used as input for the adder, where its inputs size is n bits.

The Multiplexer: The multiplexer serves to route the n-bit buses between the components in the logic blocs and the different cells in the whole matrix (more detail in the next paragraphs). The advantage of using multiplexers consists in providing high flexibility to connect n-bit buses together.

The Register: The Register is used to delay the output in order to obtain a sequential mode. The register has as inputs the LSB (Lowest Significant Bits) and MSB results (which provide from the adder or the multiplier) while its outputs are validated only with a positive edge of the clock signal.

Table 1 Logic block functions accordingly to the multiplexer's selection bits

M1	M2	M3	Function	Mode
0	0	0	A + B	Combinatorial
1	0	0		
0	1	0	(A + B) × A	
1	1	0	A × B	
0	0	1	A + B	Sequential
1	0	1		
0	1	1	(A + B) × A	
1	1	1	A × B	

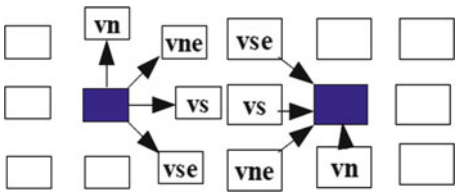

Fig. 7 Neighboring connection structure

The multiplexer M1 serves to select the first operand of the multiplier either the adder output or the second input of the logic block. The Multiplexer M2 serves to select the block LSB output between the adder output and the LSB multiplier output. The multiplexer M3 selects the function mode (sequential or combinational). The advantage of a coarse-grained architecture is the gain on configuration memory points required to select which input will be connected to the multiplexer output. With bus-based multiplexers only one SRAM bit is needed to select two n-bits buses. However, with a FGA with bit-based interconnect, n configuration RAM points are required. Table 1 describes the various functions provided by the logic block with different memory configurations.

3.2 Configurable Interconnect Architecture

As mentioned on the previous parts, the proposed architecture is a matrix of identical tiles. To elaborate the desired algorithms communication must be established between different tiles. Two levels of communication are proposed. The first one is local direct communication between the adjacent tiles and the second uses the switch blocks in order to connect non-adjacent tiles.

Local Interconnect

The adjacent connections are designed to ensure a data flow stream from bottom to top and from the left to the right as shown in Fig. 7.

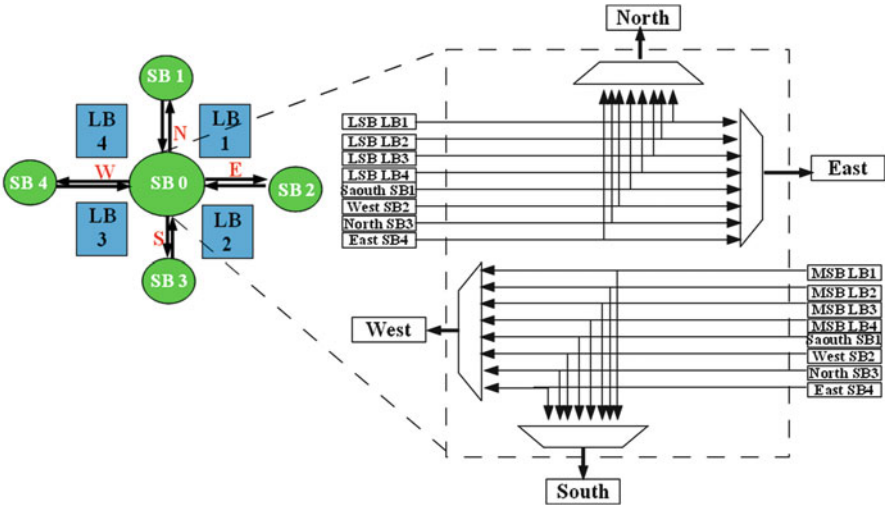

Fig. 8 The switch block architecture

Switch Blocks

Figure 8 describes the switch block topology. Its main function is to route signals of non-adjacent tiles. The number of inputs and outputs per switch box corresponds to the channel width. This parameter is very important to control architecture routability. Figure 8 shows a switch block with channel width equal to 1. To ensure a full data flow between the different rows and columns of the matrix the switch block must be able to route signals from a row to rows in both directions: top and bottom, also from a column to columns in both directions: right and the left. So, the switch block presents four outputs. These outputs are named North, South, East, and West. The North output routes the signal to the row on the top, the South output routes the signal to the row on the bottom, the East output routes the signal to the column on the right, and the West output routes the signal to the column on the on the left. As shown in Fig. 8, the switch block receives signals from two parts. The first are signals driven by adjacent switch blocks. The second are the outputs of the four adjacent logic blocks (LSB and MSB). The total number of inputs is twelve: Four inputs from adjacent switch blocks, Four LSB logic block outputs, and Four MSB logic block outputs. To choose between the 12 inputs multiplexers are needed. As four outputs are needed so the switch block will contain four multiplexers 8 to 1, a multiplexer for each output.

The advantage of using coarse-grained blocks with bus-based connections is reducing the number of memory points required to reconfigure the system.

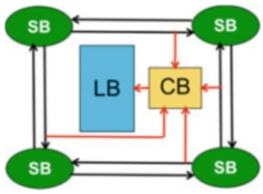

Fig. 9 Connection and switch blocks

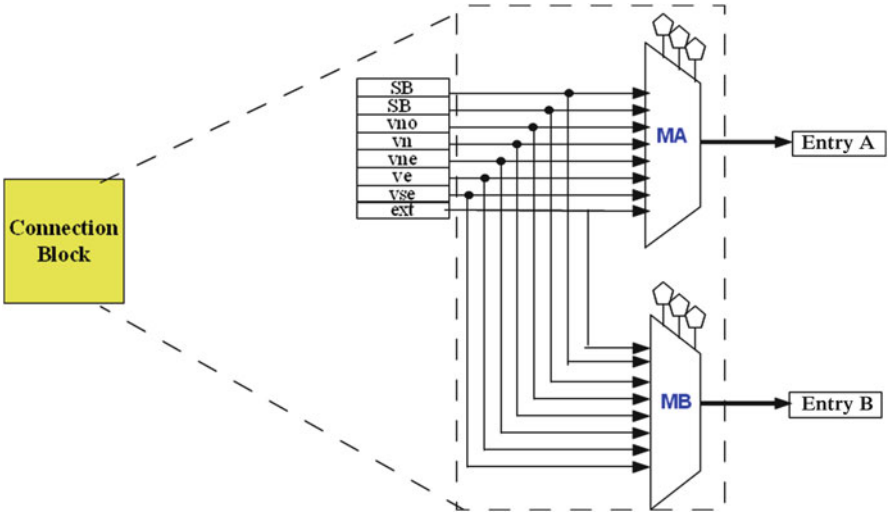

Fig. 10 The connection block topology

Connection Block

The Connection block connects the switch block outputs to the logic block inputs. It selects signals driven by adjacent logic blocks and switch block to be connected to the relevant logic block input. As discussed above there are two ways to drive signals: the direct connection between the logic blocks, only the LSB outputs, and the connection using the switch blocks. The neighboring structure described in Fig. 9 shows that four direct connections between logic blocks are considered. These direct connections represent the LSB outputs of the adjacent logic blocks. In addition to the direct connection there are other multi-bits signals driven by adjacent switch block in order to ensure data flow between different rows and columns of the matrix. Eight multi-bits signals are connected to connection block as shown in Fig. 9.

The logic block has two bus-based inputs, thus the multiplexer selects two multi-bits signals from the eight inputs using two multiplexers 8 to 1 as described in Fig. 10. The advantage of using a bus connection for data flow is the reduction of the number of memory points used to connect signals together. 2×3 memory points are only used rather than $2 \times 3 \times N$ SRAM points with fine-grained bit-based interconnect, where N is the data bus width.

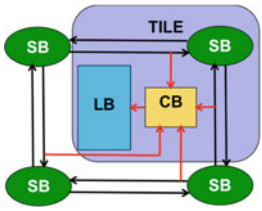

Fig. 11 Tile description: a group of logic block, connection block, and switch block

Tile Construction

The proposed architecture is a matrix of identical tiles. A tile is constituted of a logic block, a connection block, and a switch block (Fig. 11). This combination is the smallest one that can be designed and then replicated. The direct connections are not shown for figure clarity. The illustrated connections are those between the connection block and the adjacent switch blocks.

4 Architecture Exploration and Implementation

This section summarizes the realization part. First we present the adopted methodology for validation of our proposal solution. Then additional tools are performed to facilitate the generation of a VHDL description synthesizable directly by industrial tools: matrices generation and VHDL descriptions generation. Finally, we present physical layout generation and a comparison with other ASIC technology in order to show the efficiency of our method.

4.1 Architecture Exploration

To experiment the area efficiency of the proposed architecture, we place and route application netlists on the target architecture. The objective is to implement the same benchmark netlists on both fine-grained and CGAs and to evaluate the minimal required area in each case. We use ST 130 nm technology to evaluate area. To implement benchmarks on the proposed architecture we generated netlists mapped on architecture logic blocks. Each netlist is a set of instances connected with signals to achieve the following functions: DCT, FIR, FFT, etc. We developed placement and routing tools to implement various benchmarks [20]. Placement consists in placing instances communicating together in close logic blocks. In this way connections are shortened and delays are reduced. Once instances are placed, netlist signals are routed using architecture resources. In this phase routing multiplexers are configured to select which input is connected to the output. Interconnect resources of the architecture are presented by a routing graph with nodes corresponding to

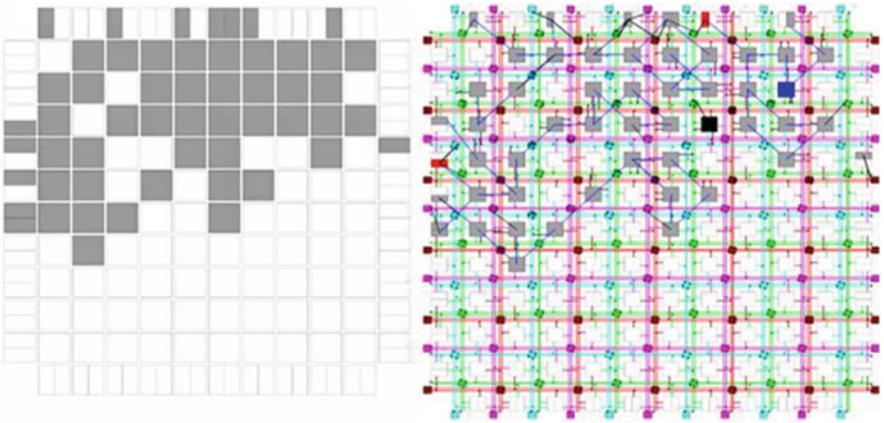

Fig. 12 Placement and routing of DCT2 A netlist (tool screen shot)

wires and CLBs/LBs pins and edges presenting switches. We use the Pathfinder routing algorithm [21], which is an iterative rip up algorithm based on the congestion negotiation. To connect terminals of each net, the router uses the Dijkstra algorithm [7] to find the shortest path (lowest total cost) between a net source node and a net sink node. At the end of an iteration, a resource can be congested because multiple nets use it. During the subsequent iterations, the cost of resources is increased. The so-called congestion cost takes into account the number of nets sharing the resource (present congestion), and the congestion history of that resource. Therefore, nets are made to negotiate for the use of routing resources.

Figure 12 shows an example of a placed and routed netlist on the proposed architecture. The white boxes correspond to architecture logic blocks. The gray boxes correspond to placed netlist instances. The utilized routing resources to route signals appear on the figure too.

Thus for each benchmark we determine the smallest matrix with minimum row, columns, and channels with capability to implement it. The obtained parameters are then given to the architecture generator to estimate the required area.

4.2 Architecture Generation

To facilitate the design and the exploration of the proposed architecture, we developed within our works, additional tools to facilitate the generation of a VHDL description synthesizable directly by industrial tools: matrices generation and VHDL descriptions generation. As shown in Fig. 13, the matrix consists of a duplication of the tile.

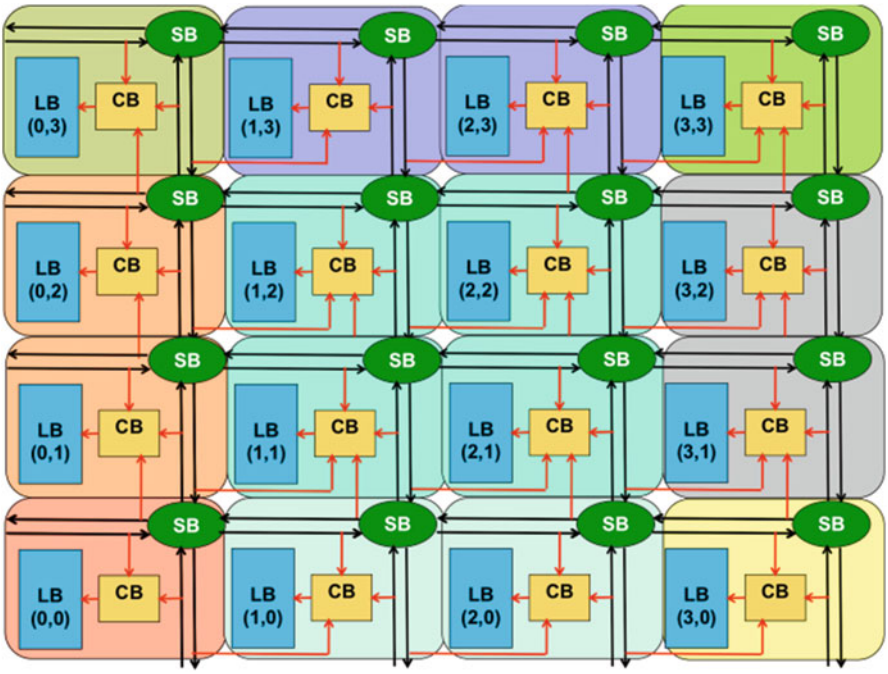

Fig. 13 Tile-based architecture generation

For the matrix generation tool, the number of tiles is equal to the multiplication of the number of rows by the number of columns. The designer specifies these parameters and the channel width. The VHDL generator is able to generate all files related to the three main blocks and evenly the files related to all components, which forms these blocks. As specified when describing the proposed blocks the size of the buses and the size of the matrix are parameters specified by the designer.

4.3 Physical Layout Generation

Figure 13 shows a 4×4 coarse-grained mesh architecture. It is divided into 4×4 tiles. Each tile is composed of a logic block, a connection box, and a switch box. We distinguish 8 different boundary tiles and 1 central tile. By duplicating these 9 tiles we can generate mesh architectures with any row and column sizes. The idea is to design a layout of each of the nine tiles and then replicate them to generate layout of any architecture size. This scalable generation technique is very beneficial to save design time and reduce recurring engineering effort. Cadence encounter is used to generate the layout of a 4×4 CGA in ST 130 nm technology.

Table 2 Characteristics of the smallest FGA implementing benchmark designs

Design	4-LUTs number	Rows × columns	Minimum channel width
DCT2A	1,273	15 × 15	14
DCT2B	1,273	15 × 15	14
FIR AT44	1,036	17 × 17	12
FIR DF44	1,036	17 × 17	12
FIR TDF44	1,036	17 × 17	12
FFT 8	506	12 × 12	10

Table 3 Design using ASIC technology, proposed coarse-grained and reference FGAs

Design	CG FPGA area (μm^2)	ASIC area (μm^2)	FG FPGA area (μm^2)
DCT2A	545,722	87,250	3,230,100
DCT2B	308,899	87,250	3,230,100
FIR AT44	422,162	74,952	2,858,000
FIR DF44	411,866	74,952	2,858,000
FIR TDF44	535,425	74,952	2,858,000
FFT 8	329,492	8,985	360,400
Average	425,594	68,056	2,565,767

4.4 Experimental Results

To validate the developed architecture several signal processing benchmarks are implemented. For each design we evaluate the smallest possible area to realize it in terms of number of rows, columns, and channel width. This exploration is run for both coarse-grained and FGAs. Since it is not possible to customize industrial FPGA to the implemented application, VPR [10] environment is used to model Xilinx Virtex architecture. As shown in (http://www-unix.ecs.umass.edu/~wxu/jbits/) [12], it is possible to model Virtex 1 and 3 architectures with VPR by using LUTs with four inputs and group them into clusters with size 4 and having 16 inputs. This architecture is used as the fine-grained reference. Table 2 shows the minimum FGA used for each design application.

According to Jonathan and Jan Kuon [15], FPGAs uses area 40 times larger than ASICs. Table 3 confirms this result and presents the area required for each algorithm using the coarse-grained technology, ASIC technology, and fine-grained technology.

Table 3 serves to measure the gap between both technologies: The proposed coarse-grained architecture gives a reduction on the gap between ASIC and reconfigurable architecture (FPGA). Thanks to unifying coarse-grained logic block and bus-based interconnections we reduce the gap between ASIC and configurable architecture area from 38 times to 7 times. Based on such large area gain we can deduce that static power consumption which is correlated to area is highly reduced compared to fine-grained-architecture. In fact, according to [13], buffers and SRAM are the major factors behind static power dissipation. As described previously using bus-based interconnects the number of required SRAM points is reduced by six

times. In addition, according to the work presented in [16], interconnect power is dominant and leakage power is significant in nanometer technologies. Thus, reducing interconnect switches leads to power consumption reduction.

5 Conclusions and Perspectives

There are clear opportunities in the design space between ASICs and microprocessors and a lot of on chip real estate value that can be added. These opportunities can be exploited by reconfigurable architectures solutions. In fact, each computational paradigm is best suited for a given range of tasks. Irregular tasks and those with complicated control flow are best solved by microprocessors when the speed of operation is not critical. When very efficient realizations in terms of both performance and power consumption are intended, ASICs are the most appropriate especially when flexibility is not an important issue. For tasks that require good performance, power consumption and flexibility reconfigurable architectures solutions present the best suited. When irregular structures and control flow dominated applications and no dynamic reconfiguration is needed FGRC solutions are convenient, but when more efficiency and dynamic reconfiguration is sought CGRC solutions are well fitted. Moreover, most DSP algorithms and their architectures are believed to be possible to implement using the CGRC solution. In fact, looking at the different DSP algorithms we note that within a given algorithm the complexity of computations and data dependencies increase dramatically within the algorithm. Hence, an accelerator implementing the complete algorithm although may need to have a larger data path can save substantial control efforts needed to partition the SFG and schedule its operation taking data dependencies in account. The selected architectures exhibited features ranging from being simple and regularity to more complex with butterfly operations to irregular structures to dynamic scheduling to memory and bit operations.

In Sect. 2 some of the CGRC solutions reported in the literature or available in the market are introduced showing their general organization, their Processing Elements (PE) architectures, their interconnect structures, and their reconfiguration characteristics. Proposed architectures aimed to narrow the gap between FPGAs and ASICs by proposing a coarse-grained FPGA architecture (CGRA) at both levels: Logic blocks or Routing connectivity. A move away from single bit operations, as are common in current FPGAs, to architectures that operate on coarser multi-bit data becomes more interesting as one way to further improve FPGA performance and efficiency. Nevertheless, most of them did not unify both coarse-grained logic and routing in the same architecture. Our solution consisted to develop an architecture shifting away from such LUT-based logic, which is inherently fine-grained, to logic units that operate on coarse-grained data and use bus-based connections to make routing. In this way we restrict our architecture to a set of signal processing applications to reduce costly flexibility and thus improve area, performance, and power consumption.

In fact, in Sect. 3 the features of the required CGRC solution are discussed in the light of the DSP architectural requirements. The proposed architecture is assimilated to a matrix, which composed by identical parts called tiles. The tile may be replicated many times as needed to construct the architecture of the algorithm. An iterative design was adopted in possessing progressivity up finding an effective solution. Three versions are proposed and the advantages and disadvantages and the degree of suitability of these architectures to the studied DSP implementations are pointed out. The general architecture is introduced and its different basic building blocks are described as well. Thus the architecture is designed corresponding to the need.

Section 4 presented two aspects of realization:

- Tools for matrices and then VHDL descriptions generation
- Experimentations and results of some design examples using the proposed techniques and comparisons with other results in order to show the efficiency of our method.

First added tools allow rapid design exploration of DSP algorithms. These tools lead also to illustration examples of several operators and DSP algorithms mapping using proposed technique. Each component is mapped using placement, routing, and layout generation industrial tools with technology ST 30 nm. In order to evaluate the results, the same designs are mapped with ASIC technology. Comparisons between the area needed on ASIC technology and implementation using proposed coarse-grained architecture are deduced. Referring to Jonathon Rose [15], the FPGAs area is 40 times bigger than the ASICs area, the proposed coarse-grained architecture gives an important reduction of the gap between configurable architecture and ASIC.

For future work, other applications such as block error correction, data compression and encryption, and the suitability proposed solutions to their realization can be studied. The above mentioned studies may result in further improvements of proposed architecture and may point out more issues that were not previously brought into attention by the studied DSP applications. Since a System on a Chip (SoC) encompassing microprocessor(s), accelerators, coprocessors, and/or reconfigurable modules, CGRC solution(s) may be used as Embedded configurable co-processor in SoC. For embedded applications, we consider more and more low consumption. We have to take into account this aspect during the design of SoC, particularly on reconfigurable modules CGRC solution some works are underway [14, 17, 26].

References

1. G. Ansaloni, P. Bonzini, L. Pozzi, EGRA: a coarse grained reconfigurable architectural template. IEEE Trans. Very Large Scale Integration Syst. **19**(6), 1062–1074 (2011)
2. M. Ayala-Rincón, C.H. Llanos, R.P. Jacobi, R.W. Hartenstein, Prototyping time- and space-efficient computations of algebraic operations over dynamically reconfigurable systems modeled by rewriting-logic. ACM Trans. Des. Autom. Electron. Syst. **11**(2), 251–281 (2006)

3. J. Becker, R. Hartenstein, M. Herz, U. Nageldinger, Parallelization in co-compilation for configurable accelerators-a host/accelerator partitioning compilation method, in *Proceedings of the Asia and South Paci_c Design Automation Conference ASP-DAC ′98*, 1998
4. P. Beshop, A tradeoff between microcontroller, DSP, FPGA and ASIC technologies. EE-Time, 25 Feb 2009. http://www.eetimes.com/
5. V. Betz, J. Rose, VPR: a new packing, placement and routing tool for FPGA research, in *International Workshop on Field Programmable Logic and Applications*, 1997
6. E. Chung, D. Burger, M. Butts, J. Gray, C. Thacker, K. Vissers, J. Wawrzynek, Reconfigurable computing in the era of post-silicon scaling [panel discussion]. FCCM, 21st annual international IEEE symposium on field-programmable custom computing machines, Seattle, 28–30 April 2013
7. T. Cormen, C. Leiserson, R. Rivest, *Introduction to Algorithms* (MIT Press, Cambridge, 1990)
8. B. De Sutter, P. Raghavan, A. Lambrechts, *Coarse-Grained Reconfigurable Array Architectures*. Handbook of Signal Processing Systems, 2nd edn. (Springer, New York, 2013), pp. 553–592
9. H. Ghariani, Design of a coarse rained FPGA architecture, Research Master, July 2011 at National School of Engineers of Sfax, University of Sfax, Tunisia
10. R. Hartenstein, M. Herz, T. Hoffmann, U. Nageldinger, KressArray Xplorer: a new CAD environment to optimize reconfigurable datapath array architectures, in *Proceedings of the Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2000
11. M. Herz, High performance memory communication architectures for coarse-grained reconfigurable computing systems, Ph.D. thesis, Universitaet Kaiserslautern, 2001
12. JBits 2.8 SDK for Virtex, Xilinx Corporation, 1999
13. S. Kaptanoglu, Power and the future FPGA architectures. in *International Conference on Field Programmable Technology*, Keynote ICFPT 2007
14. Y. Kim, R.N. Mahapatra, Dynamic context compression for low-power coarse-grained reconfigurable architecture. IEEE Trans. Very Large Scale Integration Syst. **18**(1), 15–28 (2010)
15. I. Kuon, R. Jonathan, Measuring the gap between FPGAs and ASICs. The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto
16. F. Li, Y. Lin, L. Hei, D. Chen, J. Cong, Power modeling and characteristics of field programmable gate arrays. IEEE Trans. Comp. Aided Des. **24**(11), 1712–1724 (2005)
17. L. Liu et al., An energy-efficient coarse-grained dynamically reconfigurable fabric for multiple-standard video decoding applications, in *Custom Integrated Circuits Conference (CICC)*, San Jose, 22–25 Sept. 2013
18. J. Luu, I. Kuon, P. Jamieson, T. Campbell, A. Ye, W.M. Fang, J. Rose, VPR 5.0: FPGA CAD and architecture exploration tools with single-driver routing, heterogeneity and process scaling. FPGA symposium, 2009
19. R. Maestre, M. Fernandez, R. Hermida, Bagherzadeh, F. Kurdahi, H. Sing, A framework for reconfigurable computing: task scheduling and context management. IEEE Trans. Very Large Scale Integration Syst. **9**(6), 858–873 (2001)
20. Z. Marrakchi, H. Mrabet, U. Farooq, H. Mehrez, FPGA Interconnect topologies exploration. Int. J. Reconfigurable Comput. **2009**, 1–13 (2009)
21. L. McMurchie, C. Ebeling, Pathfinder: a negotiation-based performance-driven router for FPGAs, in *Proceedings of the 1995 ACM Third International Symposium on Field Programmable Gate Arrays (FPGA 95)*, 1995
22. E. Mirsky, A. DeHon, MATRIX: a reconfigurable computing architecture with configurable instruction distribution and deployable resource, in *IEEE Symposium on FPGAs for Custom Computing Machines*, pp. 17–19 April 1996
23. A.M. Obeid, Architectural synthesis of a coarse-grained run-time-reconfigurable accelerator for DSP applications, Ph.D., University of Darmstadt, 2005
24. M. Taylor, J. Kim, J. Miller, D. Wentzlaff, F. Ghodrat, B. Greenwald, H. Hoffman, P. Johnson, J.-W. Lee, W. Lee, A. Ma, A. Saraf, M. Seneski, N. Shnidman, V. Strumpen, M. Frank, S. Amarasinghe, A. Agarwal, The raw microprocessor: a computational fabric for software circuits and general-purpose programs. IEEE Micro **22**, 25–35 (2002)

25. M. Turki, Z. Marrakchi, H. Mehrez, M. Abid, New synthesis approach of hierarchical benchmarks for hardware prototyping, in *8th International Conference on Design & Technology of Integrated Systems in Nanoscale Era (DTIS 2013)*, Abu Dhabi, 26–28 Mars 2013
26. B. Van Essen, R. Panda, A. Wood, C. Ebeling, S. Hauck, Energy-efficient specialization of functional units in a coarse-grained reconfigurable array, in *FPGA′11: Proceedings of the 19th ACM/SIGDA International Symposium on Field Programmable Gate Arrays*, 2011, 107–110
27. E. Waingold, M. Taylor, D. Srikrishna, V. Sarkar, W. Lee, V. Lee, J. Kim, M. Frank, P. Finch, R. Barua, J. Babb, S. Amarasinghe, A. Agarwal, Baring it all to software: RAW machines. IEEE Comp. **30**(9), 86–93 (1997)
28. X. Wang, et al., Mapping method of coarse-grained dynamically reconfigurable computing system-on-chip of REMUS-II, ODES ′13, in *Proceedings of the 10th Workshop on Optimizations for DSP and Embedded Systems, Shenzhen, China,* 23–24 February 2013
29. J.W. Yoon et al., I2CRF: Incremental interconnect customization for embedded reconfigurable fabrics, Design, Automation & Test in Europe Conference & Exhibition (DATE), Grenoble, 14–18 March 2011

Static and Dynamic Mapping Heuristics for Multiprocessor Systems-on-Chip

A.H. Benyamina, P. Boulet, and K. Benhaoua

Abstract In the Algerian republic, scientific and intensive computing are a priority in the strategy of the scientific research between 2013 and 2017. To put into practice this strategy, 32 nodes of high-performance computing were installed in Algeria in particular in the University of Oran. The work that we present in this chapter is our contribution to this priority in the area of intensive embedded computing: static and dynamic mapping heuristics for single chip systems.

Thus we present a synthesis of the work done over a few years in a collaboration between the University of Oran, Algeria (LIO and then LAPECI laboratories) and the DART team of LIFL (University Lille 1, CNRS, Inria), and in partnership with the Algerian Center of spatial techniques. We study how to efficiently map intensive applications on network-on-chip based multiprocessor systems-on-chip.

1 Introduction

The last years Algeria has made overwhelmed efforts to develop and increase its potentiality in intensive and embedded computing. This area constitutes the main in the strategy of the scientific research between 2013 and 2017. To concretize this strategy, 32 nodes of high-performance computing were installed in Algeria in particular in the University of Oran. To answer to this target our laboratory (LAPECI-Oran: www.lapeci.com) has developed many projects in intensive computing, real-time computing and embedded computing with the DART team from LIFL/France (Université Lille 1, CNRS, Inria) and the Algerian Center of spatial techniques as Partnership. Currently, we have obtained many results consisting of a solver for mapping complex applications on heterogenous embedded networks. We have also developed a simulator for adaptive embedded

A.H. Benyamina (✉) • K. Benhaoua
LAPECI, Univ. Oran, Oran 31 000, Algeria
e-mail: benyamina.abouelhassen@univ-oran.dz; mohammed-kamel.benhaoua@lifl.fr

P. Boulet
LIFL, UMR 8022 Univ. Lille 1, CNRS- Cité Scientifique, 59650 Villeneuve d'Ascq, France
e-mail: Pierre.Boulet@univ-lille1.fr

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_11

applications on Networks-on-Chip (NoCs). This chapter is made to introduce embedded applications and present their mapping as an important stage in the design flow as explained in the following sentences.

As silicon technology keeps scaling [5], it is becoming technically feasible to integrate entire and complex systems on the same silicon die. As a result, Systems-on-Chip (SoC) are inherently heterogeneous and therefore complex, they are often formed by multiple processors of different types (e.g., general purpose computing cores (GPP), graphical processing units (GPU), digital signal processors (DSP)) with dedicated hardware or reconfigurable logic (FPGA) and peripheral. More recently, with the increase of the number of processing cores in multiprocessor SoCs (MPSoCs), the communication hardware is becoming more complex, from hierarchies of buses to networks built from switches or routers. These architectures are thus called NoCs [7]. MPSoCs and NoCs are widely used in embedded systems where, once employed in field, they always run the same set of applications.

In this chapter we will focus on mesh-based NoC architectures, in which resources communicate with each other via a mesh of switches that route and buffer messages. A resource is generally any core: a GPP, a memory, an FPGA, a DSP. A two-dimensional mesh interconnection topology is the simplest from a layout perspective and the local interconnection between resources and switches are independent of the network size.

Nevertheless, routing in a two-dimensional mesh is easy, resulting in potential small switches, high bandwidth, short clock cycles, and overall scalability [4]. One of the most onerous tasks in this context is the topological mapping of the resources on the mesh in such a way to optimize performance (e.g., energy consumption, latency). Indeed, static mapping is a problem of quadratic assignment that is known to be NP-hard.

The search space size of the problem increases exponentially with the system size depending on the number of resources, tasks, and communications. Therefore it is one of strategic importance to define heuristic methods to search a mapping that will optimize the desired performance objectives. In addition, the strategies have to handle a multi criteria exploration of the space of possible mapping alternatives. Indeed, the objectives to be optimized are frequently multiple rather than single, and are almost always in contrast with each other (e.g., latency and energy consumption). There is therefore no single solution to the problem of exploration (i.e., a single optimal mapping) but a set of possible architectural alternatives featuring a different trade-off between the values of the objectives to be optimized (Pareto Set) [2].

A critical task for recent MPSoCs design is the minimization of the consumed energy. We start from a well-characterized task graph, a direct acyclic graph representing a functional abstraction of the application that will run on the MPSoC. Each task is characterized by the number of clock cycles used for its execution. Clearly the duration of each task and the energy spent for running it depend on the clock frequency used during the task execution but it varies in opposite ways.

The problem we face is very complex. Because we solve, at the same time, the allocation of tasks to the processors and find the optimal path allocation (or communication mapping) referred in literature as network assignment [23].

For this, we have used two methods, one based on particle swarm optimization (PSO) algorithms and the other on Dijkstra's shortest path algorithm. The solution must also verify some constraints such as area, memory, load balancing, link speed, bandwidth, and often hard real-time constraints.

The originality of our contribution is that we use a new method (PSO) in such research area and, on the other hand, we try to map tasks on processors and at the same time communications on the links of the NoC.

In the rest of the chapter we present first a static mapping heuristic, and we then propose a more recent dynamic mapping heuristic.

2 Multi-Objective Static Mapping

2.1 Problem Definition and Formulation

The application is represented by a set of tasks that communicate with each other by sending message on a heterogeneous architecture.

Definition 1. The Computation Graph (CG) is a directed graph, $G(V, E)$ with each vertex $V_i \in V$ representing a task and the directed edge (v_i, v_j), denoted as $e_{ij} \in E$, representing the communication between the tasks v_i and v_j. The weight of the edge e_{ij}, denoted by Q_{ij} represents the bandwidth of the communication from v_i to v_j.

The connectivity and link bandwidth of the NoC is represented by the NoC topology graph.

Definition 2. The NoC topology graph (NT) is a directed graph $P(U, F)$ with each vertex $u_i \in U$ representing a node in the topology and the directed edge (u_i,u_j) denoted as $f_{ij} \in F$representing a direct communication between the vertices's u_i and u_j. The weight of the edge f_{ij}, denoted by bw_{ij} represents the bandwidth available across the edge f_{ij} (see Fig. 1).

The mapping of the computation graph $G(V, E)$ on to the NoC topology graph $P(U, F)$ is defined by the one-to-one mapping function $map : V \rightarrow U$.

The mapping is defined when $|V| \geq |U|$ (see Fig. 2).

Mathematical Formulation

For a CG each node represents one task with its characteristics or property.

Let $T = t_1, t_2, \ldots, t_n$ be the set of all tasks represented by CG.

$P = p_1, p_2, \ldots, p_s$ is the set of processors represented by the nodes in NT.

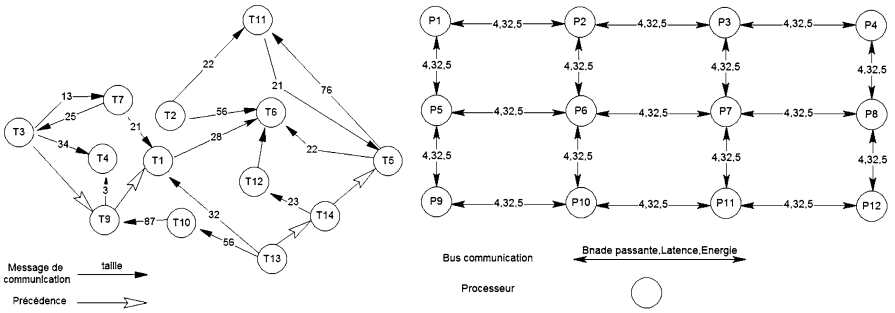

Fig. 1 Computation graph and NoC topology graph examples

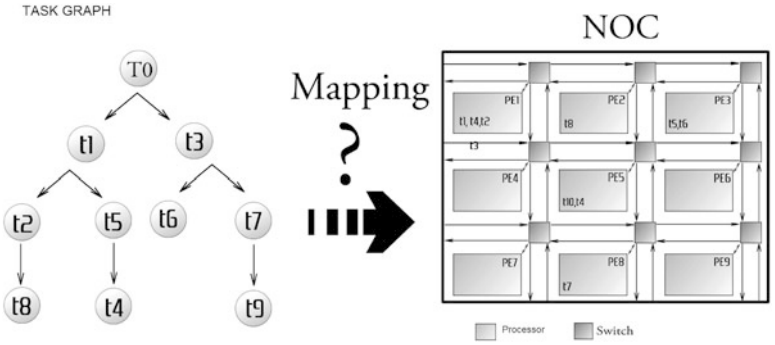

Fig. 2 Problem of mapping of a computation graph on an NoC graph

We consider that each processor p can run in different modes m_1, m_2, or m_3.

We model the allocation problem with binary variables X_{ij}^m such that [5]:

$X_{ij}^m = 1$ if $task_i$ is mapped on the processor j and runs in mode m, 0 otherwise.

$d_{ij}^m =$ duration of execution task i on processor j running at mode m.

$d_{ij}^m = \frac{WCN_{ij}^m}{f_j^m}$ where WCN_{ij}^m is the number cycle needed by $task_i$ to be executed on processor j at mode m.

f_j^m is the frequency of the clock for processor j at mode m.

dl_i is deadline for $task_i$. The time at wish $task_i$ must be terminated.

There $dl_{final} = dl_n$ is deadline of last $task_n$ and of the application.

Q_{ij} is the size of data moved between $task_i$ and $task_j$.

dQ'^m_{ijpq} is duration of communication between tasks i and j if they are assigned, respectively, to processor p and q at mode m (we see after how we compute dQ_{ijpq}^m).

q_{pq}^m is the duration of one unit (octet or bit) communication between p and q at mode m.

e_{pq}^m energy consumption for one unit from p to q at mode m.

If there is not a direct link between p and q, let $\mu(p,q) = (pi, pj)$ be path form p to q.

Then duration using links is

$$dQ'^{m}_{ijpq} = \sum Q_{ij} \times q^{m}_{plpk} \text{ where } (pl, pk) \in \mu(p,q) \text{ and } i \neq j, p \neq q. \quad (1)$$

Let us note that if tasks i and j are mapped to the same processor the duration is negligible in comparison with a case where they are mapped to different processors. Therefore in the previous expression i differs from j and also p differs from q.

And consumption for communicating $task_i$ and $task_j$ if they are assigned to p and q at mode m towards the same path is:

$$E'^{m}_{ijpq} = \sum Q_{ij} \times e^{m}{}_{plpk} \text{ where } (pl, pk) \in \mu(p,q) \quad (2)$$

Since we also take into account communication, we assume that two communicating tasks running on the same processor do not consume any energy and do not spend any time (indeed the communication time and energy spent are included in the execution time and energy), when they are allocated on two different processors, they both consume energy and spend time. Each path contains some switches and routers that need power consumption and duration [6].

Ascia et al. [2] address this problem but do not explain how to compute this. In this work we consider an average value of energy consumption (C_{sw}) and duration (d_{sw}); we can estimate these considering communications, input and output to router as stochastic.

Then if $|\mu(p,q)|$ is the length of path $\mu(p,q)$, the total consumption of switch and router on this path is:

$$|\mu(p,q)| + 1) \times C_{sw} \quad (3)$$

And the total duration is:

$$|\mu(p,q)| + 1) \times d_{sw} \quad (4)$$

We can now define the equation of the total consumption due to communication between $task_i$ and $task_j$ assigned, respectively, at processor p and q at mode m:

$$E^{m}_{ijpq} = E'^{m}_{ijpq} + (|\mu(p,q)| + 1) \times C_{sw} \quad (5)$$

and the duration due to communication as:

$$dQ^{m}_{ijpq} = dQ'^{m}_{ijpq} + (|\mu(p,q)| + 1) \times d_{sw} \quad (6)$$

We now explicit the total consumption and duration of processors.

For one $task_i$ mapped at $processor_p$ the duration is the sum of its start time and duration of all its communications with other tasks mapped to other processors.

The strategies scheduling is LS with ASAP (As Soon As Possible). Then for the $task_i$ mapped on $processor_p$ its duration D_{ip} is computed by equation

$$D_{ip} = d_{ip}^m + dQ_{ijpq}^m + (|\mu(p,q)| + 1) \times d_{sw} \text{ with } i \neq j, p \neq q. \quad (7)$$

$Dstart_i$ is the time at wish $task_i$ begin execution. It is equal at the time of the end of the last task which precedes it.

The duration of the application mapped on many processors corresponds in the time of the end of the last task. Let D be total duration including time execution and communication over links of all tasks.

$$D = Dstart_n + D_n \quad (8)$$

where D_n it is the duration of $task_n$ that is the last task. $Dstart_n$ is the time at wish $task_n$ begins execution.

The total consumption including processor, link, and switch consumption is computed by equation

$$E_{pr} = \sum_{i=1}^{N} \sum_{p=i+1}^{s} \sum_{m=ml}^{mn} x_{ip}^m \times ET_{ip}^m \quad (9)$$

where ET_{ip}^m is the consumption of $task_i$ if it is assigned on processor p at mode m. E_{pr} is thus the total energy consumed by all the processors in NoC.

$$E_{com} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \sum_{p=1}^{s} \sum_{q=p+1}^{s} \sum_{m=ml}^{mn} x_{ip}^m \times x_{jp}^m \times E_{ijpq}^m \quad (10)$$

E_{com} is the total energy consumed by network (links, routers or switch) to execute all the communications between all the tasks over all the NoC.

Multi-Objective Model for Static Mapping

Our method is based on an evolutionary computing method and a path optimization. We have to search a set of solutions *satisfying multiple objectives*. We consider here two objectives: total duration and energy consumption.

The first objective is the duration, D (8).

The second objective is the total energy consumed, E defined by

$$E = E_{pr} + E_{com} \quad (11)$$

Note that while computing D and E we look for the shortest path between two cores which satisfies the constraints (such as bandwidth and buffer size). To obtain this we have used Dijkstra's shortest path algorithm.

2.2 Mapping and Scheduling Problem Resolution

The rationale behind our approach is the minimization of the total energy consumption in order to increase the autonomy of the system. Nevertheless, trying to reach this objective increases latency. Our embedded applications are generally real time and of course the deadline must not be exceeded. Minimizing energy consumption increases latency and minimizing latency increases energy consumption. We thus have two contradictory objectives. Multi-objective problems have a set of Pareto-optimal solutions. Each solution represents a different optimal trade-off between the objectives and is said "non-dominated" since it is not possible to improve one criterion without worsening another. We propose a multi-objective approach based on the PSO technique to solve our mapping and scheduling problem.

Particle Swarm Optimization

Capone and Cesana proposed in [1] an evolutionary population-based heuristic for optimization problems. It models the dynamic movement or behavior of the particles in a search space. By sharing information across the environment over generations, the search process is accelerated and is more likely to visit potential optimal or near-optimal solutions. PSO has been extended to cope with multi-objective problems which mainly consist of determining a local best and global best position of a particle in order to obtain a front of optimal solutions. One of the well-known multi-objective techniques based on PSO algorithm is MOPSO [27]. The main algorithm is given in Algorithm 2.

The following are the phases involved in the resolution of the proposed algorithm. In continuous optimization problems, getting the initial position and velocity is more straightforward because random initialization can be used. However, since the mapping problem is a constrained optimization problem, the initial positions must represent feasible solutions. Thus, they need to be designed carefully [13].

A position in the search space represents a set of assignments that is a solution to the problem. In our case, each position provides information about how a processor in the NoC will execute each task. Then, for each position in the swarm, we assign a Boolean value to the variables X_{ij}^{m}. We consider a feasible solution, a solution that satisfies all hard and soft constraints. During the search, only non-feasible solutions that violate some soft constraints can be included in the population. This increases the likelihood of a non-feasible solution to mutate and provide a feasible one in a later generation. In the first step and as initial solution particles are chosen so that a task is allocated to a single processor on the NoC at the same time in a precise mode. For our algorithm we have tried to place the task which communicates mostly on a nearby processor or simply on same processor. To make sure that in the initial population we have some good individuals. Then for the first mapping

Algorithm 1 Algorithm for optimizing path routing

Read identifiers (processor origin Po, processor target Pt)
$Step1$ Read matrix bandwidth MB
$Step2$ Read matrix communication between tasks MC
$Step3$ Call dijkstra (Po, Pc, MB, MC)
$Step4$ Return set of links with optimal cost i.E Paths
$Step5$ Call verification (verify if bandwidth of all links still verified)
$Step6$:
if constraint not verified **then**
goto $Step2$ for another path
end if
$Step7$ Return optimal path
$Step8$ END.

we affect the task having the most neighbors on the processor having the most direct physical links. Let t_i^1 this task and p_j^1 this processor. Then $T^1 = T - t_i^1$ and $P^1 = P - p_j^1$, where T is the set of tasks and P the set of processors on the NoC. Then we redo the process for the tasks of T^1 with processors of P^1. Only here we choose in the first one task which has a link with the task t_i^1. Otherwise, we take the one which has the largest number of neighbors in absolute. We stop this initial placement when in the iteration n we obtain $T^{n-1} = T^{n-2} - t_i^{n-1} = 0$. Often the number of processors is lower among tasks ($S < N$). In this case as soon as in an any phase of this algorithm $P^k = 1$, then to initialize P^{k-1} to $P - P^k$.

Before explaining the main algorithm we have to understand how we calculate the shortest paths between two processors as well as their cost. It is the objective of the following algorithm.

Optimizing Communication and Energy Through the Network

To minimize communication and power consumption another approach is necessary to find the optimal objectives. We have used a method based on Dijkstra's shortest path (Algorithm 1). Then, we cross this method with MOPSO, described previously to approach our global objectives.

Main Algorithm Description

The entries of the algorithm are:

- An application model
- A model of the target architecture
- Constraints of latency and energy consumption
- The objective functions to optimize

Algorithm 2 MOPSO-main

input: Swarm at iteration t S^t, MaxArchiveSize, MaxIteration
Output: Repository REP
$Step0$:Initialization of Swarm
Initialize S at iteration $t = 0$
for each i $\in S^0$ **do**

  for each dimension d **do**
    Initialize $position_i$, save $pBest_i$, initialize velocity
    Specify $lowerbound_i$ and $upperbound_i$
  end for
end for
$Step1$: Evaluation of particles S
$Step2$: Update REP
for each i $\in S^t$ **do**
  compVector(i,REP)search-insert(S,REP)
end for
$Step3$: Generate Mapping(associative grid): make-Cost(Mincost)
$Step4$: Update Swarm:
for each i $\in S^t$ **do**

  for each dimension d **do**
    $Update - velocity_i, Update_position_i$
  end for
end for
$Step5$: Boundary chek
$Step6$: Update pBest
$Step7$:
if $t > MaxIteration$ **then**
  Stop
end if
$t = t + 1$ and GO TO $Step1$

And the output is an allocation of tasks and communications to hardware resources.

Our proposal for solving the problem of mapping is defined as follows.

The particle is a representation of the solution of the problem which, in this case, describes the investment. If you have an NoC mesh with S processors and an application with N tasks, then the particle is a matrix of N line and S column.

Get a set of points describing the Pareto front:

- Estimated front by iterative algorithms generate points near the front
- eliminating dominated points
- Problems of convergence:

  1. approach the front
  2. cover the entire front
  3. Concept archive: Keep each iteration all the points not dominated.

Then the problem of placement of tasks on an NoC is to minimize the objective function. It can be formulated as follows:

- A given application graph (size of the task type of Soc, runtime memory required by processor bandwidth required for a message and message size).
- A given architectural graph (speed performance by mode, power consumption by mode, load balancing (load minimum and maximum), available memory in processor, size of the queue and bus latency and energy consumption due to transmission.
- From the placement of tasks on the processor's with different modes. This is equivalent to Minimize F (X): (f (time) f (energy)). The determination of the fitness function (or adaptive function evaluation) involves several steps. Each time when a swarm is generated according to the fitness of each particle must to be evaluated.
- A particle represents a distribution of tasks of the application in the target NoC architecture.
- For a particle, move the communication costs of all messages for each message eliminating paths whose bandwidth is less than that required by the message (using the algorithm of the shortest path), then: We calculate the execution time of each task by mode which has been allocated within the processor.

At the END the Pareto Front is generated from existing solutions in the archive.

3 Multi-Objective Dynamic Mapping

Mapping of tasks onto the MPSoC platform requires finding the placement of tasks into the platform in view of some optimization criteria like reducing energy consumption, reducing total execution time, and optimizing occupancy of channels. If the MPSoC platform is heterogeneous, then a task binding process is required before finding the placement for a task. The binding process involves defining a platform resource for each task type like instruction set processors (ISPs) for software tasks and FPGA tiles for hardware tasks. Task mapping is accomplished by static (design-time) or dynamic (run-time) mapping techniques [12].

The main goal of this chapter is to present a new spiral dynamic task mapping heuristic for run-time mapping applications. The presented heuristics are applied onto NoC-based Heterogeneous MPSoC platform. Two types of PEs are considered: ISPs and reconfigurable areas (RA). Instruction set processors are used to execute software tasks and reconfigurable areas for hardware tasks. Heuristics also try to map the tasks of an application in a clustering region to reduce the communication overhead between the communicating tasks. The heuristic proposed in this paper attempts to map the tasks of an application that are most related to each other in a spiral manner and to find the best possible communication path that minimizes the communication overhead using a modified Dijkstra routing algorithm proposed

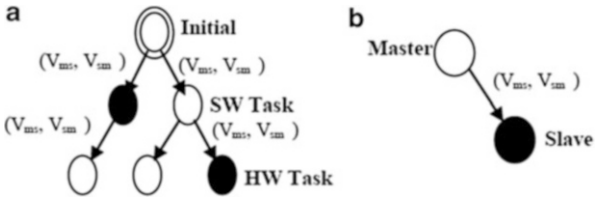

Fig. 3 Application modeling and Master-Slave

in this chapter. Our heuristic shows significant performance improvements when compared to the latest run-time mapping heuristics reported in the literature. The performance metric includes execution time and energy consumption.

3.1 Heterogeneous MPSoC Architecture

The MPSoC architecture used in this work contains a set of different processing elements which interact via a communication network [3]. Software tasks execute in ISPs and hardware tasks execute in reconfigurable logics (reconfigurable area, RA) or in dedicated IPs.

One of the processing nodes is used as the Manager Processor (M) that is responsible for task scheduling, task binding, task placement (mapping), communication routing, resource control, and reconfiguration control. The M knows only the initial tasks of the applications. The initial task of each application is started by the M and new communicating tasks are loaded into the MPSoC platform at run-time from the task memory when a communication to them is required and they are not already mapped.

3.2 Dynamic Spiral Task Mapping

We describe here our proposition in two steps. Firstly we describe our dynamic spiral task mapping. Secondly we describe the modified Dijkstra routing algorithm. But before that we introduce some definitions for a proper understanding of the proposed approach.

Definitions

Definition 3. An application task graph is represented as an acyclic directed graph $TG = (T, E)$, where T is set of all tasks of an application and E is the set of all edges in the application (Fig. 3) (a) Describes an application having initial software and hardware tasks along with the edges (E) connecting these tasks and (b) shows the master–slave pair (communicating tasks). The starting task of an application is

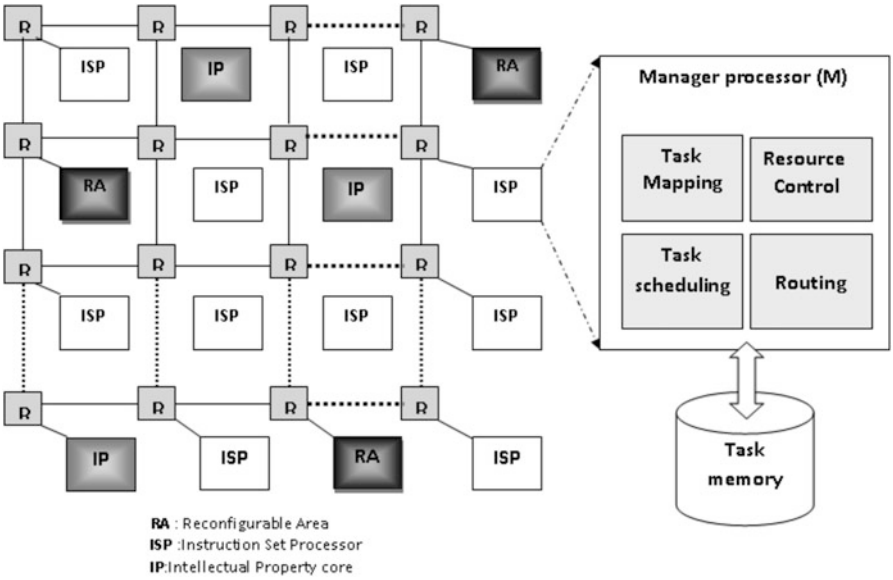

Fig. 4 Conceptual MPSOC architecture

the initial task that has no master. E contains all the pair of communicating tasks and is represented as $(mt_{id}, st_{id}, (V_{ms}, V_{sm}))$, where m_{tid} represents the master task identifier, st_{id} represents the slave task identifier; V_{ms} is the data volumes from master to slave; V_{sm} is the data volumes from slave to master.

Definition 4. An NoC-based heterogeneous MPSoC (Fig. 4) architecture is a directed graph $AG = (P, V)$, where P is the set of tiles p_i and $v_{i,j}$ presents the physical channel between two tiles p_i and p_j. A tile p_i consists of a router, a network interface, a heterogeneous processing element, local memory and a cache.

Definition 5. The application mapping is represented by a function, $map : T \rightarrow P$, allocating the tasks of the application onto the NoC-based heterogeneous MPSoC. The mapping is searched when $|T| \geq |P|$

Reference Dynamic Mappings Heuristics

1. *The First Free (FF) heuristic*: Simply selects the next compatible processor to map a given task, thus walking sequentially through all processors before considering a processor again.
2. *Minimum Maximum Channel load (MMC) heuristic*: Considers all possible mappings for a given task and chooses the one that increases the least the peak load of a channel of the NoC.

3. *Minimum Average Channel load (MAC) heuristic*: Considers all possible mappings for a given task and chooses the one that increases the least the average load of the channels of the NoC.
4. *The Nearest Neighbor (NN) heuristic*: Considers only the proximity of an available resource to execute a given task. NN starts searching for a free PE able to execute the target task near the source task. The search tests all n-hop neighbors, n varying between 1 and the NoC limits.
5. *The Path Load (PL) heuristic*: Computes the load in each channel used in the communication path. PL computes the cost of the communication path between the source task and each one of the available resources. The selected mapping is the one with minimum cost.
6. *The Best Neighbor (BN) heuristic*: Combines NN search strategy with the PL computation approach. The search method of BN is similar to NN.

Proposed Spiral Heuristic Based on Our Modified Dijikstra Routing Algorithm

1. *Spiral heuristic*: To Map the applications, firstly the initials tasks of applications are placed in distributive way the farest possible between them in a middle of the clusters, using a strategy of clusters like shown in Fig. 5. This permits the same tasks of application could be placed in a same region near between them, which reduces the communications costs. The frontiers of clusters are virtual and the common regions could be shared by the tasks of different applications.

   After the initials tasks are placed we place the most communicative tasks. To place required task, the master processor (M) tries to place it around the processor which has executed the appealed task going from a distance equal 1 (hop) until the limit of NoC. The resource (the processor according to the type of the task) is researched on spiral manner according to sequence 1, 2, 3, 4, 5, 6, 7, 8 as shown in Fig. 5. Explores spiral neighbors and performs the mapping this prevents the calculation of all possible solutions mapping, as in the PL, or calculation of the best neighbors, as in the BN heuristic, what allows to reduce the global execution time for mapping. All steps of this method are explicates in Algorithm 4.
2. *Modified dijkstra routing algorithm (Algorithm 4)*: After mapping the communicating tasks, we need a communication mapping between them. Our new proposed communication mapping tries to search a best path with a high bandwidth. The proposed heuristic reduces the computational time and energy consumption.

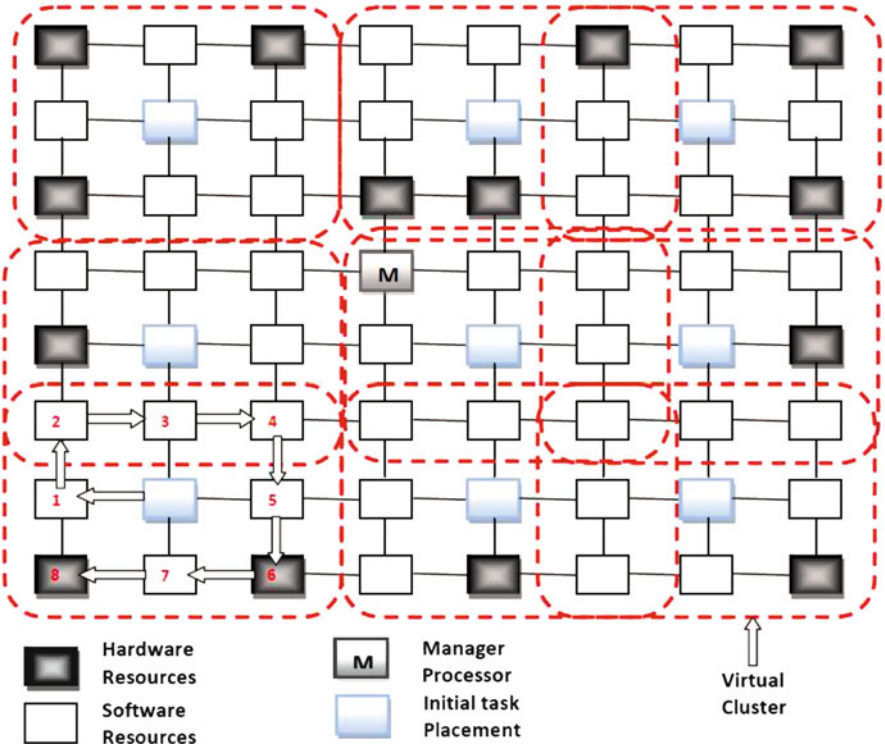

Fig. 5 Initial tasks placement for mapping applications with Spiral packing strategy

3.3 Experimental Evaluation

Experimental SetUp

We have realized a simulation of a heterogeneous platform which comprises 64 processors of which 14 hardware, 49 software, and a processor which is used like processor manager which is responsible for placement of the applications tasks, the configuration and update the platform, the communications routing. This platform uses a network on chip for the communication. We have used XML file for describing graphs of tasks used which are the same used in work of A.K. Singh, tasks (initial, software and hardware). The processing time of tasks depends on the specificity and the capacity of processor. We have fixed the parameters: software processors need 40 cycles for an instruction, however hardware processors are fast and need 20 cycles for one instruction. In a reverse of the consumption of energy or the processors hardware consume more than the processors software that we have fixed to 20 and 10, respectively. The shape of tasks is fixed to a number of instructions. The shape of exchanged data is 100 packets. The used scenario is a number of 1, 3, 7, and 10 applications which possess between 7 and 9 tasks.

Algorithm 3 Spiral packing heuristic

In: NoFreeResources, CurrentProcessor, PE, FreeResources, HopDistance, NocLimit **Out:** FreeElement

```
 1: HopDistance ← 0
 2: HopX ← 0
 3: HopY ← 0
 4: while (FreeElement = NULL) AND HopX != NocLimit AND HopY!= NocLimit do
 5:     if (CurrentProcessor.left.isFree()) then
 6:         FreeElement ← CurrentProcessor.left;
 7:     else
 8:         HopX- -;
 9:     end if
10:     if CurrentProcessor.top.isFree() then
11:         FreeElement ← CurrentProcessor.top;
12:     else
13:         HopY- -;
14:     end if
15:     if CurrentProcessor.right.isFree() then
16:         FreeElement ← CurrentProcessor.right;
17:     else
18:         HopX+ +;
19:     end if
20:     if CurrentProcessor.bottom.isFree() then
21:         FreeElement ← CurrentProcessor.bottom;
22:     else
23:         HopY+ +;
24:     end if
25: end while
```

Algorithm 4 Modified Dijkstra routing algorithm

In: IdSource, IdDestination **Out:** BestTraject

```
    /*While we did not reach the destination */
 2: while (stop=false) do
        /*Got back the neighbor who has the least used link*/
 4:     min ← minWeight();
        if min <> idDest then
 6:         /*Add the link in the list of the path*/
            Listpath.add(min)
 8:     else
            stop ← true
10:     end if
    end while
12: BestTraject ← Listpath
```

The platform is divided into nine clusters which permits to launch nine applications in parallel. Beyond this number the other applications to be placed have to wait in queue. For the placement of the tasks of the applications we have implemented our proposed dynamic task mapping based on spiral packing strategy and modified Dijkstra routing method. The spiral method tries to map the tasks in

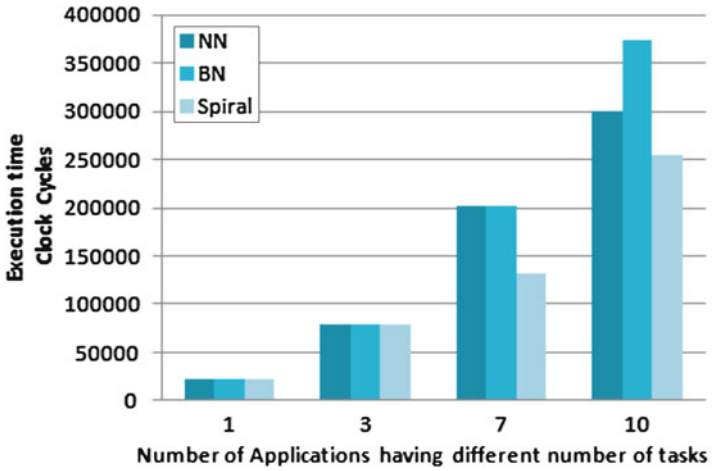

Fig. 6 Execution time comparison of the proposed approach with NN and BN, respectively

close manner with minimum exploration of the NoC space. The implementation of our method modified Dijkstra routing which minimizes the time processing and energy consumption of the system. For a comparison study we have implemented the NN and BN dynamic mapping heuristics.

Experimental Results

We have executed the implemented dynamic heuristics the Nearest Neighbor (NN) and the Best Neighbor (BN) for the placement of 1, 3, 7, and 10 applications in parallel on the simulated platform of 64 processors: 14 hardware, 49 software, and one for the processor manager (M). For the same we have executed our proposed heuristics for routing and spiral run-time task mapping. For the measurements of performances we have calculated the execution time and the energy consumption. Figure 6 shows the optimization brought by our approach in terms of execution time awards the use of the proposed approach. Figure 7 shows the optimization brought by our approach in terms of energy consumption.

4 Related Works

Most of the existing works in the literature that solve the problem of mapping on the NoC platform propose static methods. Static mapping defines task placement at design time, having a global view of the MPSoC resources and the tasks of

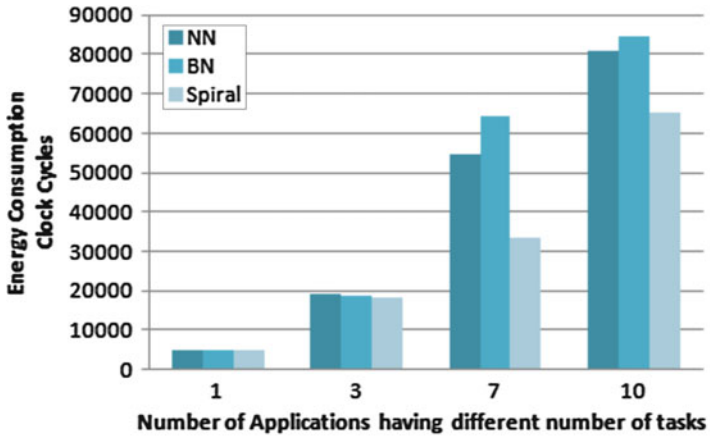

Fig. 7 Energy consumption comparison of the proposed approach with NN and BN, respectively

applications (tasks graph). As it is executed at design time, it may use complex algorithms to better explore the MPSoC resources, resulting in optimized solutions.

Heuristics like genetic approach and exact methods like tabu search and simulated annealing are presented in [11, 15, 21, 28, 29]. In [19, 23], energy-aware mapping algorithms are presented. These techniques find fixed placement of tasks at design-time with a well-known computation and communication behavior.

As it is executed at design time, static mapping may use complex algorithms to better explore the MPSoC resources, resulting in optimized solutions [12,15,19,21, 23,28,29]. However, static mapping is not able to handle a dynamic workload, new tasks or applications loaded at run-time. To cope with this feature of actual MPSoCs, dynamic (run-time) mapping techniques are required to map them onto the platform resources [9,11,16,18,20,26].

The challenge in the latest works to solve the problem of mapping in the NoC-based heterogeneous MPSoCs is to present run-time mapping techniques for mapping application's tasks onto them. Wildermann et al. [26] evaluate the benefits of using a run-time mapping heuristic (communication and neighborhood cost functions), which allows decreasing the communication overhead.

Holzenspies et al. [17] investigate another run-time spatial mapping technique, considering streaming applications mapped onto heterogeneous MPSoCs, aiming on reducing the energy consumption imposed by such application behaviors. Schranzhofer et al. [22] suggest a dynamic strategy based on pre-computed template mappings (defined at design time), which are used to define newly arriving tasks to the PEs at run-time. Carvalho at al. [10] evaluate pros and cons of using dynamic mapping heuristics (e.g., path load and best neighbor), when compared to static ones (e.g., simulated annealing and Taboo search). Carvalho's approach was extended by Singh et al. [24,25], employing a packing strategy, which minimizes the communication overhead in the same NoC-based MPSoC platform. Additionally,

Singh's approach was improved to support multitask mapping onto the same PE. Different mapping heuristics were used to evaluate the performance. According to the authors, the communication overhead of the whole system is reduced, decreasing the energy consumption. Faruque et al. [14] propose a decentralized agent-based mapping approach, targeting larger heterogeneous NoC-based MPSoCs (32×64 system is used as case study).

Mapping heuristics Nearest Neighbor (NN) and Best Neighbor (BN) presented by Carvalho and Moraes [8] and two run-time mapping heuristics presented by Singh et al. [25] have been taken for evaluation and performance in comparison with our proposed mapping heuristics.

5 Conclusion and Future Directions

In this chapter we have surveyed our contributions to the problem of static and dynamic mapping of computation intensive embedded applications on MPSoCs.

These propositions have been published in common between University of Oran, University Lille 1 and the center of spatial techniques of Algeria. We continue to work together to elaborate new solutions to the problem of mapping for more general applications (using hierarchical models) on more heterogenous networks-on-chip.

Meanwhile we also develop a middleware for adaptive applications which consider mapping and load balancing on NoCs. We hope through our works and results to participate in the effort granted in Algeria in the field of research and the application software in the economy and the industry.

References

1. E. Amaldi, A. Capone, M. Cesana, I. Filippini, F. Malucelli, Optimization models and methods for planning wireless mesh networks. Comput. Net. **52**(11), 2159–2171 (2008)
2. G. Ascia, V. Catania, M. Palesi, A multi-objective genetic approach to mapping problem on network-on-chip. J. Univers. Comput. Sci. **12**(4), 370–394 (2006)
3. L. Benini, G. De Micheli, Networks on chips: a new soc paradigm. IEEE Comput. **35**(1), 70–78 (2002)
4. L. Benini, D. Bertozzi, A. Guerri, M. Milano, Allocation, scheduling and voltage scaling on energy aware mpsocs, in *CPAIOR* (2006), pp. 44–58
5. A.E.H. Benyamina, P. Boulet, Multi-objective mapping for noc architectures. J. Digit. Inf. Manag. **5**(6), 378–384 (2007)
6. A.E.H. Benyamina, P. Boulet, B. Beldjilali, An hybrid algorithm for mapping on noc architectures, in *META'2008* (2008)
7. T. Bjerregaard, S. Mahadevan, Survey of research and practices of network-on-chip. ACM Comput. Surv. **38**(1), 1–51 (2006)
8. E. Carvalho, F. Moraes, Congestion-aware task mapping in heterogeneous mpsocs, in *SOC* (2008), pp. 1–4

9. E. Carvalho, C. Marcon, N. Calazans, F. Moraes, Evaluation of static and dynamic task mapping algorithms in noc-based mpsocs, in *SOC* (2009), pp. 87–90
10. E. Carvalho, C. Marcon, N. Calazans, F. Moraes, Dynamic task mapping for mpsocs. IEEE Des. Test Comput. **27**(5), 26–35 (2010)
11. C.-L. Chou, R. Marculescu, User-aware dynamic task allocation in networks-on-chip, in *Proceedings of DATE* (2008), pp. 1232–1237
12. C.-L. Chou, C. Mellon, R. Marculescu, Incremental run-time application mapping for homogeneous nocs with multiple voltage levels, in *Proceedings of Hardware/Software Codesign and System Synthesis* (2007), pp. 161–166
13. H.J. Escalante, M. Montes, L.E. Sucar, Particle swarm model selection. J. Mach. Learn. Res. **10**, 405–440 (2009)
14. M.A.A. Faruque, R. Krist, J. Henkel, Adam: run-time agent-based distributed application mapping for on-chip communication, in *Proceedings of DAC* (2008), pp. 760–765
15. P. Ghosh, A. Sen, A. Hall, Energy efficient application mapping to noc processing elements operating at multiple voltage levels, in *NOCS* (2009), pp. 80–85
16. P.K.F. Hölzenspies, G.J.M. Smit, J. Kuper, Mapping streaming applications on a reconfigurable mpsoc platform at run-time, in *Proceedings of SOC* (2007), pp. 1–4
17. P.K.F. Hölzenspies, J.L. Hurink, J. Kuper, G.J.M. Smit, Run-time spatial mapping of streaming applications to a heterogeneous multi-processor system-on-chip (mpsoc), in *Proceedings of DATE* (2008), pp. 212–217
18. M. Mandelli, A. Amory, L. Ost, F.G. Moraes, Multi-task dynamic mapping onto noc-based mpsocs, in *SBCCI* (2011)
19. A. Mehran, S. Saeidi, A. Khademzadeh, A. Afzali-Kusha, Spiral: a heuristic mapping algorithm for network on chip. IEICE Electron. Express **4**, 478–484 (2007)
20. A. Mehran, A. Khademzadeh, S. Saeidi, Dsm: a heuristic dynamic spiral mapping algorithm for network on chip. IEICE Electron. Express **5**(13), 464–471 (2008)
21. S. Saeidi, A. Khademzadeh, F. Vardi, Crinkle: a heuristic mapping algorithm for network on chip. IEICE Electron. Express **6**(24), 1737–1744 (2009)
22. A. Schranzhofer, J.-J. Chen, L. Santinelli, L. Thiele, Dynamic and adaptive allocation of applications on mpsoc platforms, in *ASP-DAC* (2010), pp. 885–890
23. D. Shin, J. Kim, Power-aware communication optimization for networks-on-chips with voltage scalable links, in *Proceedings of CODES and ISSS* (2004), pp. 170–175
24. A.K. Singh, W. Jigang, A. Prakash, T. Srikanthan, Efficient heuristics for minimizing communication overhead in noc-based heterogeneous mpsoc platforms, in *Proceedings of International Symposium on Rapid System Prototyping* (2009), pp. 55–60
25. A.K. Singh, T. Srikanthan, A. Kumar, W. Jigang, Communication-aware heuristics for run-time task mapping on noc-based mpsoc platforms. J. Syst. Archit. **56**(7), 242–255 (2010)
26. L.T. Smit, J.L. Hurink, G.J.M. Smit, Run-time mapping of applications to a heterogeneous soc, in *Proceedings of SOC* (2005)
27. D.-L. Truong, Shared protection for multi-domain networks. Ph.D. thesis, University of Montreal, 2007
28. X. Wang, M. Yang, Y. Jiang, P. Liu, Power-aware mapping for network-on-chip architectures under bandwidth and latency constraints. TACO 2010 **7**(1), 6 (2010)
29. Y. Zhang, X. (Sharon) Hu, D.Z. Chen, Task scheduling and voltage selection for energy minimization, in *Proceedings of Design Automation Conference* (2002), pp. 183–188

Part IV
On New Governance Models

Using Mobile Phones in Elections in Developing Countries: Opportunities and Challenges

Rehema Baguma and Emmanuel Eilu

Abstract A considerable number of developing countries have introduced Information and Communication Technologies (ICTs) mainly in the initial stages of the electoral process such as voter registration using biometric technology. Post-conflict and emerging democracies in particular are exploring biometric voter registration to have clean voter registers—free of ghost voters such as Democratic Republic of Congo, Togo, Guinea Conakry, and Uganda. However, several other attempts to integrate ICTs—particularly computer-based systems in electoral processes in many developing countries have registered high failure. Causes of failure are mainly social cultural and economic challenges rather than technical and security factors. The various failures come at a soaring price for the world's poorest countries. For example in 2001, Uganda lost US$ 22 million in a failed e-registration system. Hence there is a need to continue investigating how such failures can be avoided or at least reduced. Some Scholars have suggested the use of mobile phones as a voting tool given their wide spread use, portability, and affordability. However, to-date, research on use of mobile phones in elections has mainly focused on individual cases of using mobile phones in elections in specific countries. To our knowledge, there has not been an examination of current use of mobile phones in elections versus the opportunities they offer and existing challenges. This discussion is useful to inform future development of mobile phone based electoral services and enactment of relevant policies. This chapter discusses opportunities mobile phones can offer developing countries in organizing and conducting free and fair elections, how mobile phones are currently being used, challenges still inhibiting wider use of mobile phones in elections in developing countries and possible solutions.

R. Baguma (✉) • E. Eilu
School of Computing and Informatics Technology, Makerere University,
P.O. Box 7062, Kampala, Uganda
e-mail: rbaguma@cis.mak.ac.ug; rehema.baguma@gmail.com; eiluemma@cis.mak.ac.ug

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_12

1 Introduction

Election is a process through which a group of people express their opinion over who will lead them for a specific period of time [28]. Olaniyan et al. [28] note that key requirements for the integrity of an electoral process include; correctness, robustness against fraudulent behaviors, and coherence.

The United Nation's Millennium Declaration on democracy states that democracy and participatory governance must be based on the will of the people. For the last 20 years, competitive free and fair elections have become the standard in some African countries. In nations such as Ghana, Mauritius, and South Africa, free and fair elections have not only helped consolidate emerging democratic institutions but also improved projection for greater economic and political development. In other countries like Liberia and Sierra Leone, plausible elections have shaped the route for national reconciliation and a return to democratic rule after years of armed conflict and civil war. On the other hand, blemish elections are a recipe for tragedy and aggression especially in many African countries where elections are characterized by manipulation of the rules governing elections through enacting legal but unfair rules and manipulation of voter choice through bribery, rigging, intimidation, and violence. This often sparks off post-election crises such as the case in Kenya in 2007, in Zimbabwe in 2008, in Uganda in 2011, and in Nigeria in 2011 [12]. Flawed elections in countries such as Kenya in 2007 and Zimbabwe in 2008 led to violence, loss of life, and destruction of property and have further polarized political discourse. As a result, voter turn-up is declining in most developing democracies with the African continent having the lowest voter turn-up of 65 % [34]. This calls for continuous efforts to find better ways to organize and conduct elections to facilitate free and fair voting processes and to improve voter turn-up. In Sub-Saharan Africa, electoral violence and intimidation, vote-buying, and ballot-fraud are rampant [2]. Collier and Vicente [11] noted that these irregularities have been used strategically by politicians to bend electoral outcomes. The following subsection discusses how ICTs are being used in elections in developing countries.

1.1 Use of ICTs in Elections in Developing Countries

There is increasing use of Information and Communication Technology (ICT) to facilitate free and fair electoral processes due to the realization that ICT has the potential not only to improve transparency, privacy, efficiency, and effectiveness of electoral processes but also to improve the interaction between citizens and their governments through various e-participation systems. Based on this premise, a number of ICT aided electoral systems have emerged over years and these have in some cases aided in detecting fraudulent practices and have improved voter turnout and transparency especially in developed countries and some developing countries like Brazil and India.

Common types of e-voting include; polling station where voters cast their votes electronically on an electronic machine within the polling booth; kiosk e-voting where voters cast their votes at pre-selected stations through ATM-like terminals; and remote e-voting where voters cast their votes anywhere and anytime there is Internet access as well as voting through mobile devices [28].

A considerable number of developing countries have introduced ICTs mainly in the initial stages of the electoral process such as voter registration using biometric technology. Post-conflict and emerging democracies in particular are exploring biometric voter registration to have clean voter registers—free of ghost voters. Examples of countries that have so far carried-out biometric registration in Africa include; Democratic Republic of Congo, Togo, Guinea Conakry, Uganda, Angola, Nigeria, and Mozambique [8]. However, several other attempts to integrate ICTs—particularly computer-based systems in electoral processes in many developing countries have registered high failure [15, 19]. Examples of failure cases include; the failed electronic voter registration systems in East Timor, Kosovo and Nigeria in 2007, Peru in 2000, Nicaragua in 2008, and Uganda in 2001 among others [7, 15, 23]. According to [7], technical complications and high costs have created significant obstacles for successful implementation of sophisticated electronic procedures in new, fragile, and transitional democracies. Olaniyan et al. [28] attributes failure in ICT-aided electoral systems to lack of proper infrastructure, low information literacy, low Internet penetration rate, lack of skilled personnel, and expensive specialized electronic voting hardware among others. Heeks [15] attributes the failures to the design reality gap. This view is shared by [21, 23, 26, 35] who note that social cultural and economic challenges rather than technical and security factors are the major causes of failure in ICT-aided elections in developing countries. For example, in terms of cost, ICT-aided electoral systems come at a soaring price for the world's poorest countries yet most of these governments depend on donor funding to finance these initiatives. On the social-cultural side, ICT electoral systems do not match the realities of societies/governments in that they do not fit into the organizational, human, social, and cultural issues [15].

The various failures come at a high price for the world's poorest countries. For example in 2001, Uganda lost US$ 22 million in a failed e-registration system [28]. Hence there is a need to continue investigating how such failures can be avoided or at least reduced. Some Scholars have suggested the use of mobile phones as a voting tool given the wide coverage of mobile phone networks [8]. Compared to the Internet, mobile phones exhibit unique characteristics that distinguish them from the Internet namely: portable, affordable, and in wide spread use [30].

Globally, mobile-cellular penetration rates stand at 96, 128 % in developed countries, and 89 % in developing countries [20]. In Uganda, by June 2011, the mobile phone subscription rate had reached 14 million [31]. According to Yuijuico [38], 80 % of the Philippine population have access to a mobile phone and 1.39 billion messages are sent daily compared to an estimated 24 million (26.7 %) who use the Internet. One of the main attractions of text messaging as opposed to voice calls is the low cost. Yuijuico [38] notes that unlike in developed countries where a minute of airtime and a text message cost nearly the same, the least expensive

prepaid airtime in the Philippines costs roughly $0.10 a minute while a text message costs $0.028. In Uganda, a minute of a phone call costs roughly $0.08 whereas sending short messages (SMS) costs roughly $0.02.

Building a voting service on a widely used platform like this has cheaper and better potential than one that is less widely used and less familiar to the electorate.

1.2 Mobile Phone Services in Elections

Mobile phone services that can be used in elections include:

SMS to mobile phones of supporters: With mobile phone text messaging, one can send an abstract or light version of content (up to 160 characters) or call to action on the small mobile phone screen, and refer to more detailed content usually via web or e-mail. Text messaging is a major factor behind the popularity of mobile phones in addition to its voice capabilities mainly due to the low cost compared to voice services. Unlike in developed countries where a minute of airtime and a text message cost nearly the same, a call in the Philippines costs roughly $0.10 a minute while a text message costs $0.028 [38]. In Uganda, a minute of a phone call costs roughly $0.08 whereas an SMS costs roughly $0.02.

Vote casting by SMS: SMS voting is just another way of sending a vote to the recipient by using a short code instead of using a normal mobile number. SMS voting provides a useful, secure, and reliable process as mobile users are informed of the charges incurred and receive a return SMS to acknowledge that their SMS vote has been received [29].

Combining Mobile technology and social media: While mobile phone and SMS technology might have greater reach now, many entrepreneurs say the potential lies in integrating social media and mobile applications as election monitoring tools [36]. Citivox, a startup using crowd-sourcing technology in Mexico to enhance civic engagement and transparency, use integrated reporting channels that include SMS and social media sources when collecting reports from people on elections and other fronts. The aim of Citivox's approach is to close the feedback loop based on the fact that one cannot separate civic life from social life.

Other services include:

- Politically themed ringtones for callers to hear or download into their phones.
- Use of a short code to recruit supporters for rallies and other campaign events.
- Raising money from supporters by charging instant donations to their phone bill or mobile money for countries where it exists and
- Encouraging supporters to forward text messages, ring tones, and short codes to friends and family.

2 Related Work

To-date, mobile phones are being used extensively in a range of social, economic, and political activities in developing countries. For example, for monitoring measles outbreak in Zambia, supporting diagnosis and treatment by health workers in Mozambique and sending health education messages in Benin, Malawi, and Uganda [3]. In Kenya, Malawi, and South Africa, mobile phones are being used to send several reminders daily to HIV-positive patients [3]. Farmers in Niger, Senegal, Ghana, Kenya, Uganda, and India access M-agricultural services such as market prices and extension services via the mobile phone [3]. In addition, mobile phones are also being used in political campaigns such as the Save Mabira forest campaign and a campaign to boycott sugar produced by the company that the government wanted to give the forest in Uganda in 2008 [16], the Nairobi People's Settlement Network used mobile phones to get organized against evictions in Kibera slum, Nairobi [16] and mobilization for protests against the president of Philippines which eventually led to his ousting [38]. Other forms of mobile phone data services in use in developing countries include; sending money to family members and friends [5], buying items from the market, banks notifying customers when a withdrawal or deposit has been made or during promotional offers, paying utility bills such as electricity, water, etc.

Besides the social, economic, and general political mobilization, there have also been attempts in a couple of developing countries to use mobile phones in elections. However, to-date, research on use of mobile phones in elections has mainly focused on individual cases of using mobile phones in elections in specific countries. Key cases in point include; sending women and youth informational SMS alerts on voting and registration procedures and asking for their feedback via text on the electoral process in Egypt in 2013 [36]; taking stock of residents' views in the lead up to the 2012 poll in Libya [36]; SMS reminders to vote in Thailand [38]; campaign slogans in form of ring tones, wall paper, and SMS reminders to vote by the Bharatiya Janata Party in India in 2004 [22]; sending text messages about what has been witnessed to a data center for specialists to process and map in 2011 in Egypt [36]; election monitoring, reporting human rights abuses, strengthening civil society, and democratizing the flow of voting information in Indonesia in 2005, Palestine in 2006, Bahrain in 2006, Albania in 2007, Sierra Leone in 2007, and Lebanon in 2009 by the National Democratic Institute [6]; using a hotline for citizens to SMS incidents of electoral fraud and voting irregularities in Tunisia in 2011 [36]; reporting irregularities or fraud in Sierra Leone's general elections in 2007 [4]; voters' call for attention when it appeared that votes for the Coalition for Unity and Democracy (CUD) party were being stolen at polling stations [9]; large scale election monitoring for the Montenegrin Referendum on Independence in 2006 [38]; mobilization of tens of thousands of young demonstrators in the streets of Kiev in 2004 to protest election fraud and

demand a revote [4]; use of mapping data tools in Afghanistan's 2009 presidential elections for easier presentation of data [36]; verification of voter registration in Kenya in 2007 [16]; SMS reminder to potential voters for the referendum in 2005 by the Uganda Electoral Commission [16]; a violence-prevention tool for mobile monitoring and reporting of post-election violence in Kenya [16]; Bunge SMS-Civil and political mobilization services for empowering Kenyan citizens to influence local governance in their constituencies [16].

To our knowledge, there has not been a cross cutting examination of current use of mobile phones in elections versus the opportunities they offer and existing challenges. This discussion is useful to inform future development of mobile phone based electoral services and enactment of relevant policies.

This chapter discusses opportunities mobile phones can offer developing countries in organizing and conducting free and fair elections, how mobile phones are currently being used, challenges still inhibiting wider use of mobile phones in elections in developing countries and possible solutions. The rest of this paper is organized as follows: opportunities mobile phones can offer developing countries in organizing and conducting elections and how they are currently being used; challenges inhibiting wider use of mobile phones in elections in developing countries and possible solutions.

3 Opportunities Mobile Phones Can Offer Developing Countries in Elections and How They Are Currently Being Used

A number of scholars have noted that mobile phone voting if adopted well, promises a much cheaper, and more secure voting process than paper ballots, far simpler vote counting and high voter turn-up compared to other voting technologies. Additionally, given the current sophistication of mobile devices with audio, video, web, and e-mail, individual users and networks can send and receive polling data with custom software, view partisan candidate messages, and hear politically motivated ringtones. Stein [30] noted that mobile phones are more pervasive than Internet access, so they offer a means to reach a broad constituency of citizens during an electoral period. Across the world, mobile phones are being used to register voters, educate citizens on specific candidates or referenda, track electoral activities, fraud, and raise money for candidates and parties [30]. Stein [30] and Heeks [29] noted that young people that are often the target of voter registration campaigns are very comfortable using mobile phones and text messaging in particular. This section discusses opportunities mobile phones can offer in the organization and execution of elections and how they are currently being used in elections in developing countries.

3.1 Voter Education

Mobile phones can be used to boost voter awareness at a significantly low cost around elections through SMS. In 2013, in Egypt, Souktel targeted women and youth, sending them informational SMS alerts on the voting and registration procedures, and asking for their feedback via text on the electoral process [36]. In Libya, the Souktel Company worked with Al Jazeera to take stock of residents' views in the lead up to the 2012 poll and the project received approximately 5,000 responses [36]. In 2006, the Thai Election Commission sent messages to 25 million mobile phone customers reminding them to vote [38]. According to Yuijuico [38], another campaign message advised voters to bring a pen and paper with them to mark their ballots to prevent fakes. Stein [30] noted that the already common use of text messaging in Philippines was an important factor in the success of use of text messaging in voter education in Philippines in 2006. According to Yuijuico [38], 1.39 billion text messages are sent daily in Philippines and this has earned the country the title, 'the world's text messaging capital'.

3.2 Campaigning

Mobile phones are a highly personal means of communication that allow campaigners to bypass the mass media when targeting voters with electoral messages [30]. According to Stein [30], political ringtones, wallpapers, and SMS election updates are one of the effective ways to campaign in elections. In India's 2004 general elections, the Bharatiya Janata Party set up a team to generate campaign slogans to be transmitted via mobile phones and this helped them reach better urban voters who may have otherwise been apathetic [22].

3.3 Candidate and Political Party Support/Mobilization

Mobile phones also bring tangible benefits to campaigns of political candidates and parties as these explore ways to mobilize supporters and voters. Phones allow for targeted messaging to supporters that have opted in to receive such communications. In addition, the viral nature of forwarding text messages has led to a smart mob effect whereby political supporters keep each other informed of developments as they evolve in real time [30]. Yujuico [38] noted that young people are an active driver of this cultural evolution who can spontaneously organize youth meet-ups using short messaging service. Throughout the world including in developing countries, there is already use of mobile phones/sms in other forms of campaigns/mobilization efforts such as fundraising by NGOs, individuals organizing social/family functions like weddings, funerals, etc. Stein [30] noted that mobile phone based fundraising is cheaper compared to other communication means such as print, radio, etc.

3.4 Election Monitoring

This can be achieved through employing SMS and voice-based reporting schemes to keep an eye on polling stations and undertake public opinion polls. Although social media networks like Facebook are often credited with aiding the Arab world's revolutions, mobile phones remain a more powerful tool in monitoring elections given their reach especially in Sub-Saharan Africa where Internet penetration and literacy rates are low. In 2011, Project Rakeeb partnered with Advanced Computer Systems Firm to screen Egypt's first post-revolution parliamentary elections. Observers at polling stations used their cell phones to send text messages about what they witnessed to a data center, where specialists then processed and mapped the reports [36]. The National Democratic Institute successfully deployed mobile phone technology to improve election monitoring, report human rights abuses, strengthen civil society and democratize the flow of voting information in countries such as Indonesia in 2005, Palestine in 2006, Bahrain in 2006, Albania in 2007, Sierra Leone in 2007, Lebanon in 2009, and many others [6]. When Tunisia held its first national elections in 2011, after uprisings ousted former President Ben Ali, Souktel partnered with the Tunisian Bar Association to set up a hotline for citizens to SMS incidents of electoral fraud and voting irregularities in which a team of 80 lawyers received roughly 1,000 reports in a customized web platform [36]. In Sierra Leone's national election held in 2007, about 500 observers were posted at polling stations around the country to report on irregularities or fraud using mobile phones and at the end, both local and foreign observers declared it to be free, fair, and transparent based upon which Sierra Leone is now used as an example for other African countries to follow [4]. In the May 2005 parliamentary elections in Ethiopia, voters used their mobile phones to call the attention of the CUD party when it appeared that their votes were being stolen at polling stations which prevented fraud [9]. Yujuico [38] reported that in 2006, mobile phones were used in large scale election monitoring for the first time during the Montenegrin Referendum on Independence.

3.5 Post-Election Protests

Howard [17] in his book "*Smart Mobs*" describes how effective the proliferation of mobile phones can bring together groups at a moment's notice to protest. He believes that mobile phones are catalyst for change and that anonymity of users makes these protests unusually difficult for police or other authorities to stop. Citizens or Smart mobs, brought together by text messages, have led to political change in countries like the Philippines and Ukraine. In January 2001, opposition forces in Philippines mobilized an estimated 250,000 people using mobile phone SMS to match on Epifano de los Santos Avenue (EDSA) to oust then-President Joseph "Erap" Estrada over corruption allegations [38]. Estrada's successor, Gloria Macapagal-Arroyo took over for the remaining 3 years and later contested in

2004 and subsequently won but controversy dogged this result when wiretapped phone calls of her discussing vote tallies with the Election Commissioner Virgilio Garcillano were released in 2005 [38]. According to [38], following this revelation, Philippines was filled with cell phone pop culture and ringtone remixes featuring Arroyo asking Garcillano "*Hello Garci? So will I still lead by 1 M?*" —close to the nationwide margin of victory. Yujuico [38] noted that Arroyo's popularity never returned to pre-"*Hello Garci?*" levels, although she managed to retain her presidency up to the end of her term. In 2004, SMS messages in Ukraine helped mobilize tens of thousands of young demonstrators in the streets of Kiev to protest election fraud and demand a revote in what was termed as the Orange Revolution [4]. Similarly in 2005, bus drivers in Ethiopia sent messages to each other calling for a "stay-at-home" protest that lasted for 2 days crippling the economy of the country [4].

3.6 Increased Levels of Accuracy

It is scientifically agreed that no two thumb-prints are the same. The mobile voting system could automatically invalidate any individual who attempts to vote twice thereby rendering such votes void. Hence, electoral integrity would be upheld by ensuring that an individual is entitled to only one vote. Trying to verify votes manually could be an overwhelming exercise. The voter verifiable paper audit trail of the mobile voting system would serve as evidence of an election that took place. If properly programmed, the mobile voting system would provide information such as the period of commencement and termination of the elections as well as the individuals who voted as well as the time and other details [18]. In cases of dispute, the thumb print, representing cast votes of individuals who were delimited to a particular phone could be compared to their earlier votes to ascertain their authenticity.

3.7 Easy to Consume Data Presentation

Data is an essential national infrastructure just as important as the roads and telecommunications infrastructure. Once the job of collecting and analyzing data is done, how the information is presented is crucial in making it relevant. Various visualization tools are helping election monitors and officials to better present their findings. In elections where SMS has been used, one can watch data flow into the database directly when it is time for the monitors to report. The system automatically sends confirmation messages back to the observer in an interactive exchange of SMS messages, so accuracy increases. At reporting time, numbers simply change on the screen as the SMS messages pour into the database [4]. According to Cheng [9], using tools like MapBox, an open source software, SMS voting data can

be visualized by layering different strains of information onto custom maps, from vote counts at polling stations down to details about literacy and population. The maps can then be published online. This kind of stacking of various data sets makes identifying patterns that are not otherwise obvious possible. Development Seed applied mapping data tools in Afghanistan's 2009 presidential elections and there is hope that in the 2014 presidential elections, open data work is going to help domestic observers, international observers, the government, civil society organizations, and the media to deliver that data more efficiently to the public [36].

3.8 Increasing Accessibility for People with Disabilities

When mobile phones are used in elections/voting, people with disabilities gain a new alternative for voting privately, independently, and with dignity and at their own pace. This can be achieved through use of online/mobile tools which they are familiar with such as customized screen readers, specialized input devices, etc. According to National Union of Disabled persons in Uganda (NUDIPU) [27], there are over 16 million people with disabilities in Uganda and one million of these are blind. Such a population would benefit a lot from mobile phone based electoral services.

4 Challenges Currently Faced or Hindering Use of Mobile Phones in Elections & Possible Solutions

4.1 Security

In countries that still have low levels of democracy like is the case in many sub-Saharan African countries, it is not safe to openly oppose government or monitor and report its non-democratic activities. Therefore security of people participating in anti-government voting activities is important for such activities to succeed. Stein [30] noted that there are significant security concerns for some activists using mobile phones during elections in areas where SIM cards and accounts are registered and monitored by local authorities as is the case in Uganda. In Ethiopia in June 2005, SMS was blocked during post-election protests in June after bus drivers sent messages to each other calling for a "stay-at-home" protest that lasted for 2 days crippling the country's economy [4]. A possible solution to ensuring security of participants is to hide the identity of participants. For example when Souktel partnered with the Tunisian Bar Association to set up a hotline for citizens to SMS incidents of electoral fraud and voting irregularities, only a person's mobile number was available in order to guard their identities. This was done to give participants confidence that they will not be persecuted for reporting an incident of electoral

fraud. This made participants safe and protected [36]. In response to hackers trying to hit the system of Citivox during monitoring of Mexico's midterm elections, details on citizens or monitors sending in election reports were encrypted [36].

The other aspect of security is ensuring the Electoral body or other government agencies do not compromise systems in favor of one party against the others. This is only possible with independent electoral commissions which is difficult to have in countries where Electoral Commissions are appointed by the executive.

4.2 General and Technology Illiteracy

SMS and cell phone technology as a mobilization, voting, and election monitoring medium is dependent on users being able to read and write. UNESCO [33] reported that one of the challenges of delivering mobile phone based services is that 41 % of the population in developing countries is non-literate and even the literate among the poor are typically novice users of computer technologies. According to Uganda Demographics Health Survey (UDHS) [32], the literacy rate of 15–24-year olds in Uganda is 76.1 %. UDHS defines literacy as the 15–24-year olds who have attended secondary school or higher or who can read a whole sentence or part of a sentence. White [37] defines technological literacy as the ability to understand or familiarity with common computer/mobile interaction paradigms such as the concept of menu systems, radio buttons, navigation, or data entry. Chipcase [10] observed that non-literate populations avoid complex functions and primarily use phones for synchronous voice communication. Medhi et al. [24] revealed that non-literate and novice users of technology face several barriers with existing text-based mobile interfaces such as difficulties understanding hierarchical structures, soft keys, scroll bars, non-numeric inputs, and specialized terminology. Wharton Technology [36] noted that use of SMS succeeded in Tunisia and Lebanon because the number of illiterate people is relatively low but in some developing countries, there are still millions of people who cannot read and write, and usually the great percentage are women.

To address challenges of general and technology illiteracy, there is need for tools that people can use with their voices such as Freedom Fone, Zimbabwe's IVR system [14] Freedom Fone, does not require the Internet or smartphones, but instead relies on SIM cards and conventional phone service to get running. Additionally, Souktel also offers voice-based services for example they provided an audio library service they provided to Somali citizens, who could call a hotline and learn more about civic subjects such as the constitution-drafting process. Other advantages of voice as opposed to data interfaces is that voice is compelling and harder to fake when conveying reports and content can be easily transferred to public outlets like radio broadcasts. However, Medhi et al. [24] noted that it is not an easy medium to manage.

4.3 Divided Efforts

Many organizations undertake election monitoring individually, but streamlining and merging efforts can make them more influential. When Citivox started out in Mexico's 2009 midterm elections, they found about a dozen groups with similar aims and they were able to align them on the same high-tech reporting platform to track elections. Independent groups using such technology are, in a way, providing a service that some governments are unable to perform, because their operations lack that level of sophistication. During Mexico's 2009 midterm elections, the department of administrative modernization was found doing analytics on a dry-erase marker board, managing reports with stacks of file folders and carrying out "real-time mapping" with a map pasted to a wall [36].

4.4 Poor User Experience

Despite the potential mobile phones have over other ICTs in increasing voter turn-up, its acceptance and usage as an end device for voting in developing countries has been a challenge. In Uganda for example, close to 60 % of the population are not in favor of using their mobile phones in voting while in Malaysia about 50 % are also not in favor [13]. Whereas in Malaysia, poor user experience of using mobile phones in voting is mainly due to socio-cultural factors [29], in Uganda, the main causes are political [13]. According to a study carried out by [13] in Uganda, trust in the government as a whole and the electoral commission in particular to conduct free and fair mobile phone aided voting was lacking among the population. Eilu and Baguma [13] provide eight persuasive and acceptance design strategies that can increase the acceptance rate of using mobile phones for voting in a developing country like Uganda. These include; effective computing, tailoring, social and political learning, reminder/suggestions, tunneling, convenience and proximity, perceived ease of use, and perceived usefulness.

4.5 Lack of Supporting Legislation

In 2006 in Nigeria, the then chief of the Electoral Commission secured executive approval but later faced the litmus test of convincing the legislature of the need to adopt e-voting for subsequent Nigerian elections. The National Assembly voted against the adoption of e-voting systems (including mobile phone voting) citing the immaturity of the Nigerian electorate as well as arguing that the process is prone to manipulation. For this reason, Section 53 of the Nigerian Electoral Act reads: "Voting at an election under this Act shall be by open secret ballot; the use of Electronic voting Machine for the time being is prohibited" [18].

4.6 Some Opportunities Still Difficult to Exploit for Developing Countries

Some opportunities of using mobile phones in voting are still difficult to attain for developing countries. For example whereas online donations have made an impact in US presidential elections, the same is difficult to achieve in developing countries. Yujuico [38] notes that in contrast to a developed country like the United States of America where a vast majority of citizens have credit cards or PayPal accounts, this is not the case in the Philippines where use of financial products is still not widespread. The alternative for developing countries is mobile money which is gaining rapid acceptance and use in many developing countries across the world. The number of registered mobile money users in Uganda were 2.9 m in 2011 [25].

4.7 Stage of Political Maturity

It is evident that mobile phones in elections as a tool for campaigning, awareness, monitoring and protesting is very vital for any country to have free, fair and transparent elections. However, Akoh [4] noted that for most developing countries to benefit from use of mobile phones and other technology tools in elections, there must be political will by all stakeholders pushing for fairness and transparency. This stage is yet to be achieved in a number of developing countries especially in Sub Saharan Africa.

Other challenges include:

- Mobile phones have yet to reach some rural areas in some countries due to low returns on investment in such areas. According to Hellström [16], promotion of mobile penetration and increased accessibility in areas that are not commercially viable is a big challenge to implementation of m-governance. As a solution, governments have designed incentive schemes such as universal funds and Rural Communication funds to attract investment in such areas but these are not efficient enough.
- Mobile phones offer limited interactivity, very small screens, short messages and complicated commands. Advanced phones with bigger, colour screens that are GPRS enabled are still too expensive for users in developing countries [16].
- Mobile campaigners need to be cautious about sending excessive amounts of SMS to opt-in participants. A high volume of SMSs may overwhelm participants and cause them to resent and abandon the cause prematurely.

5 Conclusion

This chapter has discussed opportunities mobile phones can offer developing countries in organizing and conducting free and fair elections and how they are currently being used namely; voter education, campaigning, candidate and political party support/mobilization, election monitoring, post-election protests, increased levels of accuracy, easy to consume data presentation and increasing accessibility for people with disabilities. It has also examined challenges still inhibiting wider use of mobile phones in elections in developing countries viz. security, general and technology illiteracy, divided efforts, poor user experience**,** lack of supporting legislation, some opportunities still difficult to exploit for developing countries, stage of political maturity, mobile phones yet to reach some rural areas in some countries, etc. It has also suggested solutions to the current challenges such as hiding the identity of participants through encryption, use of voice-based tools instead of text tools, streamline and merge efforts of organizations involved in e-electoral activities, use user experience principles when developing mobile based electoral services to improve their user experience, enact relevant legislation, etc. We hope this will benefit future development of mobile phone based electoral services and enactment of relevant policies.

References

1. C.J. Aker, P. Collier, C.P. Vicente, Voter education using cell phones: evidence from a field experiment in Mozambique, 2013
2. C.J. Aker, Is information power? Using mobile phones and free newspapers during an election in Mozambique, 2013
3. C.J. Aker, I.M. Mbiti, Mobile phones and economic development in Africa (2010), http://www.cgdev.org/files/1424175_file_Aker_Mobile_wp211_FINAL.pdf
4. H.A. Akoh, Mobile phones: transforming the electoral process. Special prize of the jury at the PIWA "Information Societies" prize, 2008
5. R. Baguma, Mobile money services in Uganda: design gaps and recommendations. HCI **2013**, 249–258 (2013)
6. B. Bardall, Election violence monitoring and the use of new communication technologies, in *Democrats, Dictators, and Demonstrators Conference Issue*, vol. 7, issue 2, 2010
7. J. Blanc, Challenging the norms and standards of election administration *Electronic Voting* (2007), http://www.ifes.org/publication/de870bce4b39339d89a3469c8de3a0a0/1%20IFES%20Challenging%20Election%20Norms%20and%20Standards%20WP%20ELVOT.pdf
8. S. Caarls, The use of technologies in electoral systems prepared For Ifes-Egypt (2010), http://www.aswatna-eg.net/assets/files/Technology%20and%20Elections.pdf. Accessed Nov 2011
9. V. Cheng, Online & mobile voting: accessibility for all (2014), http://www.haayo.org/Mobile-phones-Transforming-the.html, http://www.openideo.com/challenge/voting/concepting/online-mobile-voting-increase-accessibility-for-all. Accessed 12 Feb 2014
10. J. Chipchase, Understanding non-literacy as a barrier to mobile phone communication (2005), http://www.kiwanja.net/database/document/report_literacy_barrier.pdf. Accessed 16 May 2012
11. P. Collier, C.P. Vicente, Violence, bribery, and fraud: the political economy of elections in Sub-Saharan Africa draft, 2010

12. Colloquium on African Elections (CAE) Best practices and cross-sectoral collaboration final report (2009), http://www.ndi.org/files/African_Elections_Best_Practices_ENG.pdf
13. E. Eilu, R. Baguma, Designing a reality fit M-voting, in *7th International Conference on Theory and Practice of Electronic Governance*, Seoul, 22–25 Oct 2013
14. Freedom Fone Lab (2012), https://dev.freedomfone.org/
15. R. Heeks, Most e-government-for-development projects fail: How can risks be reduced? (2003), http://unpan1.un.org/intradoc/groups/public/documents/cafrad/unpan011226.pdf
16. J. Hellström, Mobile phones for good governance – challenges and way forward, 2010
17. R. Howard, Smart mobs: the next social revolution, 2002
18. S. Igbolade, Nigerian elections 2007: E-Voting as a missing link. Article produced in the framework of PIWA "Information Societies" prize (2008), http://www.haayo.org/Nigerian-Elections-2007-E-Voting.html
19. IT Fact 85% of electronic voting projects failed in developing countries: Fact recorded on November 5, 2004. Categories: E-government (2004), http://www.itfacts.biz/85-of-electronic-voting-projects-failed-in-developing-countries/1942. Accessed 24 June 2011
20. ITU. The world in 2013: ICT Facts & Figures (2013), http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf. Accessed 10 Feb 2013
21. Kailasam, R. m-Governance ...Leveraging Mobile Technology to extend the reach of e-Governance, 2010
22. A. Kumar, A.K. Srivastava, Designing and developing secure protocol for mobile voting, 2006
23. R. Lopez-Pintor, Assessing electoral fraud in new democracies: A basic framework. IPSA-ECPR Joint Conference, North-South University of Sao Paulo, Brazil, 2011
24. I. Medhi, S. Patnaik, E. Brunskill, S.N.N. Gautama, W. Thies, K. Toyama, Designing mobile interfaces for novice and low-literacy users. ACM Trans. Comput. Hum. Interact. **18**(1), Article 2 (2011)
25. A. Mugenyi, Wanted: a mobile money regulator. The CEO Magazine, June 2012
26. M. Muthoni, Using JAD to bridge the design-reality gaps; A major cause of IS projects' failures in the developing countries, in *The Proceedings of the 2nd Annual International Conference on Sustainable ICT in Developing Countries*, Makerere University, Kampala, 6–9 August 2006
27. National Union of Disabled Person in Uganda (NUDIPU), http://www.nudipu.org/
28. O.M. Olaniyan, T. Mapayi, S.A. Adejumo, A proposed multiple scan biometric-based system for electronic voting. Afr. J. Comput. ICT **4**(2), 9–16 (2011). ISSN 2006–178
29. N.K. Sin, A. Sulaiman, A.H.S. Zolait, Analysis of mobile users' perception towards SMS voting. Commun. IBIMA **1**(21), 182–190 (2008)
30. M. Stein, Strategy guide: using mobile phones in electoral and voter registration campaigns, 2006
31. Tentena, Mobile phone subscribers hit 14 million in Uganda. Business week, 20th June (2012), http://www.busiweek.com. Accessed 10 Sep 2012
32. Uganda Demographic Health Survey (UDHS) (2011), http://www.ubos.org/onlinefiles/uploads/ubos/UDHS/UDHS2011.pdf. Accessed 20 June 2012
33. UNESCO Institute of Statistics Data Centre, regional literacy rates, 2007
34. C. Vergne, Turnout in developing countries: the effect of mass media on national voter participation. Center for studies and international development CERDI, 2009
35. G. Walsham, D. Robey, S. Sahay, Foreword: special issue on information systems in developing countries. MIS Q. **31**(2), 317–326 (2007)
36. Wharton Technology: Using Mobile Phone Technology to Ensure Fair Voting in the Middle East (2013), http://knowledge.wharton.upenn.edu/article/using-mobile-phone-technology-to-ensure-fair-voting-in-the-middle-east/
37. G. White, Designing Mobile Services for non-literate communities, in *Proceedings of 2nd International Conference on M4D-Mobile Communication Technology for Development*, Kampala, 2010
38. E. Yujuico, Democratisation & new voter mobilization in Southeast Asia: digital democracy and voter mobilisation in the Philippines, 2010

e-Government Service Delivery Model (eGovSDM): Aspects of Namibia's Road-Map to Vision 2030

Mbale Jameson, Maria Ntinda, and Stefanus van Staden

Abstract In the mid twenty-first century, attempts by some sub-Saharan nations to set Goals and Visions to attain successful service delivery to their citizens were made but only to be hampered by numerous barriers, such as the lack of infrastructure, funding, trained manpower and resources. Other barriers were poor planning, management, regulatory laws, bureaucratic processes and accountability. Namibia was not an exception to such barriers that negatively impacted on service delivery to its citizens. In view of the service delivery challenges faced, the e-Government Service Delivery Model (eGovSDM) was envisaged to move service delivery to higher levels of access and performance that will support the achievement of Namibia's Visions 2030. The eGovSDM consisted of eleven functional components in this work classified as elements or strategies and these were: Approach and Principles, Financing, Monitoring and Evaluation, Infrastructure/Platforms, Access, Capability, Content, Interaction, Collaborations, Processes and e-Applications. These eGovSDM interlinked elements or strategies were focused on solving individual targeted service delivery barriers to attain enhanced Service Delivery Performance.

1 Introduction

e-Government Service Delivery Model (eGovSDM) is envisaged to facilitate the implementation of Namibia's Road-Map to achieve Vision 2030. The eGovSDM (see Fig. 1) demonstrates the eleven (11) interlinked components in this work classified as elements or strategies that drives service delivery performance. These elements or strategies are: Approach and Principles, Financing, Monitoring and Evaluations, Infrastructure/Platforms, Access, Capability, Content, Interaction,

M. Jameson (✉) • M. Ntinda
Department of Computer Science, Centre of Excellence for Telecommunications, University of Namibia, Windhoek, Namibia
e-mail: mbalej@yahoo.com; jmbale@unam.na; pewantinda@gmail.com

S. van Staden
Office of the Prime Minister, Government Buildings, Windhoek, Namibia
e-mail: sstaden@opm.gov.na

A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, DOI 10.1007/978-3-319-08239-4_13

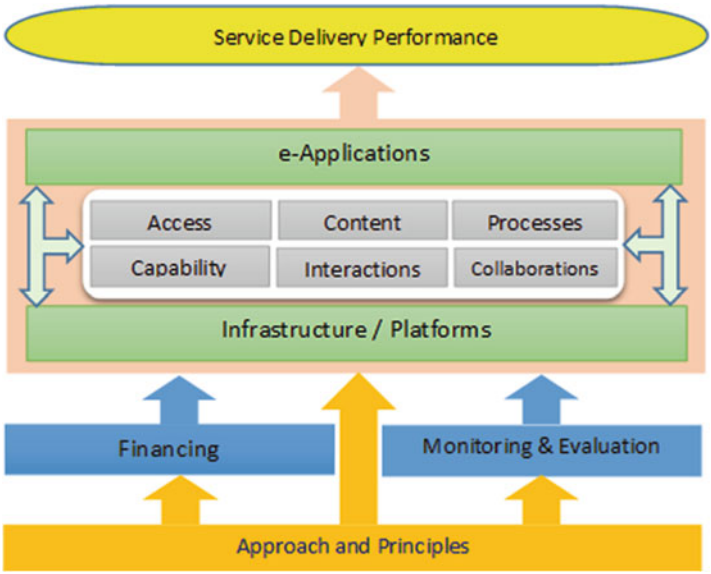

Fig. 1 e-Government enhanced service delivery model

Collaborations, Processes and e-Applications. The individual elements would be used to solve the targeted barriers of service delivery such as cumbersome bureaucracy, lengthy processes, slow service delivery, high costs of infrastructure and maintenance, high transaction costs, lack of content, lack of Information and communication technology (ICT) legislation and policies, lack of trained manpower, but to mention a few. The effective implementation of the eGovSDM would provide adequate solutions to combat the barriers and stimulate Service Delivery Performance of public and private sectors. In this manner, the eGovSDM would facilitate the public sector (i.e., Government) to address the indicators of good governance such as the efficiency, accountability, citizen participation and equity. The addressing of these indicators will lead Namibia to sustained success in achieving its Vision 2030 objectives. Another consequence is that the Namibian Government would be able to improve the management of service delivery across all Government sectors.

1.1 Statement of the Problem

Attempts have been made by various sub-Saharan countries to implement e-Government services, but with limited success due to a number of barriers encountered in the process. Some of the major barriers are: IT/ICT infrastructure, organisational issues, personnel skills, operational costs, privacy and security.

Zakareya and Zahir [28] also identified and discussed a number of barriers restricting the implementation of e-Government infrastructure, which prevented the realisation of benefits and these included, IT infrastructure, security and privacy, IT skills, organisational issues, and lack of effective leadership and operational costs. The Third National Development Plan [25] Report also discussed the major barriers to Namibia's transition into a knowledge-based and technology driven nation and these were: lack of adequate numbers of qualified personnel, the low disposable household income and lack of access to electricity in rural areas, which limits access to ICT, lack of a reading culture, which resulted in the low literacy rate.

The above identified and discussed barriers are not exceptional to Namibia, for it also finds some difficulties to implement service delivery to its citizen. It also faces the challenges of capacity, IT infrastructure, organisational and integration issues, security and costs. All these slow down the service delivery across all walks-of-life. In order to overcome the discussed barriers, a comprehensive model with specific elements or strategies targeting each barrier needed to be designed to simultaneously provide immediate solutions. It is in view of this that the eGovSDM is envisaged to improve the service delivery performance in the private and Government sectors. The eGovSDM is designed and developed equipped with a series of different application elements or strategies that would be individually targeted to specific barriers. Each element or strategies are structured to solve particular problems or barriers.

1.2 Objectives of eGovSDM

The eGovSDM is developed in line with the following objectives:

- Design the eGovSDM that would enhance Service Delivery Performance across all Government sectors.
- Formulate the eGovSDM's elements or strategies such as: Approach and Principles, Financing, Monitoring and Evaluation, Infrastructure, Access, Capability, Content, Interaction, Collaborations, Processes and e-Applications that would stimulate the Service Delivery Performance.
- Automate all the elements or strategies of the eGovSDM.
- Envisage a Road-Map to facilitate the attainment of Vision 2030.
- Create a portal or one-stop-shop counter 24 h a day, 7 days a week where interaction among Government, private and citizens would be done.
- Cut down the complex and over stretched bureaucratic system.
- Decentralise responsibilities and Processes where each citizen could contact the Government through a website where all forms, legislation, news and information are made available.
- Enforce all line-Ministries to adapt XML as the standard for data integration and browser for interface.

- Write manuals that could be used by communities from all walks-of-life for the deployed infrastructure.
- Design and deploy easy interface to be comfortably used by the illiterate communities.
- Adapt the Gartner (2000) e-Government Maturity Model which provided different e-Government solutions to suit citizen choices.
- Design systems where complete transactions could be done online, making citizens not queuing.
- Integrate all forms of Processes, allowing the citizens enjoy the services of G2C and G2B.
- Harmonise legal framework for e-Government of Namibia in line with International trends to ensure better service delivery to citizens.
- Strategise and continued budgetary support across all Offices, Ministries, Agencies (OMAs) of the Namibian Government.
- Adapt the eGovSDM institutionalised Approach where citizens from all walks-of-life would be involved with the e-Government Processes rather than a championed-led one.
- Utilise the eGovSDM Principle to drive all the elements or strategies to enhance the Service Delivery Performance and bring the Namibian Government closer to citizens and businesses.
- Envisage the Principles of infrastructure that were critical to the Namibian e-Government requirements for the attainment of Vision 2030.
- Equip all the Namibian citizens with high level relevant skills across various technologies in order to keep pace with trends of new developments.
- Build the tailor made Content translated into the local Namibian languages to give an equal opportunity even to the illiterate citizens to confidently read the materials.
- Encourage all line-Ministries to use e-Applications across all walks-of-life.
- Address the indicators of good governance in service delivery processes such as efficiency, accountability, participation and equity.
- Establish strong and continuous funding schemes with the public and private sectors in order to sustain the programmes service delivery in the Government.

1.3 Organisation of the Chapter

The chapter is organised in the following parts: Sect. 1 introduces the eGovSDM, highlighting the components and its functions in the model. Section 2 discusses similar models done by other researchers presented in two categories, those from outside the country and the others locally. The eGovSDM model is presented in Sect. 3, by articulating the role each component play to facilitate the Service Delivery Performance. Section 4 gives the analysis on how significant some components are to the overall function of the eGovSDM. Section 5 summarises the functions and benefits of the eGovSDM model.

2 Literature Review

The concept of e-Government had been discussed at various forum and platforms. In this work, the authors' first look at how it has been dealt else was outside Namibia and then locally.

2.1 e-Government External Experiences

Odile [16] explained that ICT had helped to deliver more modern services for citizens and businesses, stimulate the Information Society and emerging new economy, drive public sector transformation and help governments prepare for future pressure on public administrations. She stressed that ICT has become an integrated part of everyday life, and the next generation of e-Government will have to continue to improve public sector performance. She also pointed out that what was next is the need for a change of focus and direction: making administrations efficient and effective by putting e-Government services online was no longer enough. She highlighted four themes for a future agenda emerge that included: increasing coherency and integration of the public sector through innovation and change; putting users at the steering wheel of the public sector; allowing front-line personnel to deliver public services efficiently and effectively and globalisation of public services.

Stephen [23] described e-Governance to be the convergence between the two trends: resulting in new ways of governing via new information and communication channels. He emphasised that e-government has the potential to improve the performance of public institutions and make them more transparent and responsive; facilitate strategic connections in Government by creating joined-up administrations in which users could access information and services via portals or "one-stop-shops"; and empower civil-society organisations (CSOs) and citizens by making knowledge and other resources more directly accessible. He further outlined ten major administrative and democratic improvement prospects e-Governance offers such as: cheaper and more effective management and processing of information; a freer flow of information between departments, agencies and layers within government; more professional administrators, supported by standardised, electronically embedded decision-making systems; but to mention a few.

Stephen [23] and Heeks [11] strongly pointed out the barriers of e-Governance in Africa, claimed that 85 % of e-Government projects in developing/transitional countries were partial or total failures. In fact, Berman and Tettey [3] supported by stressing that the success rate of introduced information technology systems in African state agencies had been distressingly low, and the capacity-building objectives remained largely unachieved. They further said that African e-Governance faces two additional barriers: the lack of ICT infrastructure and mass connectivity to the internet, and the existence of post-colonial administrative cultures dominated by under-resourced and unaccountable bureaucracies.

Subhajit [24] pointed out that it was not difficult for people in developed countries to imagine a situation in which all interaction with government could be done through one counter 24 h a day, 7 days a week, without waiting in lines. He then argued that to achieve this same level of efficiency and flexibility for developing countries was going to be far more difficult. He further mentioned that experience in developed countries showed that this would be possible if Governments were willing to decentralise responsibilities and processes, and if they started to use electronic means. He went on to emphasise that each citizen could then contact the Government through a website where all forms, legislation, news and other information made available. He then wished and assumed that the Governments of developing countries, as a collector and source of information, could follow this trend, to serve its customers (citizens and businesses) better and to save costs by making internal operations more efficient, cutting down the complex and over stretched bureaucratic system.

Subhajit [24] summarised e-Government as when it has reached the stage where Departments collaborated in significant ways to avoid duplication of efforts, and a one-stop contact point was created, which was capable of handling procedures of all involved Departments. When reached this level, all information systems were integrated and the public could get G2C and G2B services at one (virtual) counter, that's one single point of contact for all services would be the ultimate goal.

Backus [2] and Subhajit [24] explained that e-Governance was more than just a government website on the Internet. They highlighted that the strategic objective of e-Governance was to support and simplify governance for all parties, government, citizens and businesses. They also stressed that the use of ICTs could connect all three parties and support processes and activities. Their emphasis was that, in e-Governance electronic means support and stimulate good governance. They added that the objectives of e-Governance were similar to the objectives of good governance and good governance could be seen as an exercise of economic, political, and administrative authority to better manage affairs of a country at all levels.

Economist Intelligence Unit [8] discussed e-Government centred on benefits and trend. The report outlined the following benefits: getting the most out of public funds, Boosting computer literacy, encouraging citizen participation, and investing in infrastructure and delivery. Also the Report discussed the trends: growing demand for transparency and accountability, targeting corruption, connecting the back-end, enhancing service delivery, closing the e-Government divide, offering multi-channel service delivery and beyond m-Government.

2.2 e-Government Local (Namibia's) Experiences and Efforts

As the world-wide attempts were made establish and implement e-Government services, Namibia was not exceptional, in this twenty-first century also had been striving to adopt it. The Office of the Prime Minister (OPM) [17] described the

efforts of the country in establishing and implementing e-Government. The OPM Report stated that the Government had now begun to recognise the needs beyond efficiencies, service delivery and web presences, to consider a more broad vision of values, ethics, culture and the needs of an integrated harmonious society, with the view to improve innovativeness in the entire range of operations conducted. The Report argued that when e-Governance was applied properly yielded the following benefits: (1) e-Governance drove costs of service delivery down and improved efficiency, (2) it was possible to do more work, with less people, in less time with a smaller workforce, (3) e-Governance also brought suppliers closer to its customers, and this increased the willingness of customers to proactively involved in Government matters, beyond just being recipient of services, (4) the Government was made to seek ways that would make it more responsive to the needs of both the citizens and the business with which they interacted, (5) ICT provided an almost limitless number of ways in which services could be provided and improved, (6) Government showed accountability and transparency, and espoused the values of decentralisation and marketisation.

The OPM [17] also narrated additional efforts the Government undertook to develop e-Governance by establishing some units such as: (a) the Department of Public Service Information Technology Management (DPSITM) was created to guide and oversee all aspects of ICT usages in the Public Service and it worked to develop open standards and cooperative architecture; (b) the Cabinet Committee on IT (CCIT) was established to ensure that ICT matters were considered at the topmost decision levels of the country; (c) the Public Service Committee on IT (PSCOIT) was created to draw membership from all Government units and its main stakeholders, to oversee and advise on the formulation of strategies that will stimulate the effective use of ICT in the government, in a manner that will ensure efficient decision-making, better service delivery and improvement in all Government management processes. OPM gave a detailed illustration of DPSITM by highlighting that it was formed to establish common standards and infrastructure to enable interoperability across Government Ministries. The OPM further emphasised that the standard would also ensure that the Government Ministries could communicate electronically with citizens and businesses.

The OPM [17] reported that the Government had already developed an Intranet that linked many Government Ministries. It (OPM) summed up its discussion by reporting that the Government of Namibia intended to provide high quality e-Governance services that would lead to increased productivity and competitiveness in the private sector by reducing the cost of the public service and the cost at the business side, increasing efficiency and reliability of information and service provisions.

eGov Namibia Report [9] narrated that the Namibia Government conducted an e-Readiness assessment and the findings from International benchmarking had provided a good insight in terms of interventions required at various areas. In view of that, the Report informed that the vision for e-Government was formed which stated: citizen and business friendly, transparent, affordable and accessible delivery of information and services to all through a professional, responsive and networked Government.

The eGov Namibia Report [9] summed up by highlighting the key Critical Success factors that would influence the commencement and implementation Government included the following: first, coming on board of all the OMAs in full spirit of the EGSAP; second, continued budgetary support, across all OMAs, there must be continued political and executive commitment which was a pre-requisite for successful implementation; but to mention a few.

3 e-Government Service Delivery Model

Much of the efforts attempt to develop and implement e-Government by the Namibian Government had been intensively and extensively discussed in the previous sections. In view of that, this work analysed the previous discussion and envisaged the eGovSDM model shown in Fig. 1, which would be suitable to enhance future development through the utilisation of elements and strategies, e-Government benefits, trends and solutions to improve the service delivery in the Government and private sectors.

The envisaged eGovSDM as already discussed above, consisted of eleven functional components, in this work they are referred to as elements or strategies and these are: Approach and Principles, Financing, Monitoring and Evaluation, Infrastructure/Platforms, Access, Capability, Content, Interaction, Collaborations, Processes, e-Applications.

3.1 Approach

The planned systematic Approach in this case would lead the eGovSDM acquisition of road-map to the attainment of Vision 2030. At the root of the model is the Approach and Principles.

The well planned approach towards the development of the e-Government is a critical element to address all the factors required for a successful implementation of such projects. In this work, it is emphasised to adopt an institutionalised Approach, where all citizen from all walks-of-life are involved with the e-Government processes rather than a champion-led one (Approach). According to Arc.gov [1] he illustrated that there had to be a step-wise approach to e-Governance so that outcomes are maximised and citizens reap early benefits from e-Governance. Therefore, from the envisaged model, eGovSDM, a sum total of well planned: e-Applications, Collaboration, Infrastructure, Processes, Content and Financing would lead to a greater and effective service delivery performance of e-Government system. Hence, a clear and flexible approach would always help to integrate all the infrastructure components of that particular model, in this case the eGovSDM. Hernán [12] explained that flexible Approach also allowed for groups of Departments and department-specific infrastructures to interconnect with the common

infrastructure as appropriate. From his explanation, it was clear that eGovSDM components would interconnect with all common elements or strategies of the model to effectively deliver the service appropriately. In fact, the Approach in this work serve as a tool used to gain strong foundation based utilising the critical planning at the initial stage.

3.2 Principles

Core guiding Principles as indicated in the eGovSDM are essential for the success of e-Governance. The Principles of the eGovSDM guide the model's elements or strategies such as: e-Applications, Infrastructure, Access, Capability, Content, Interaction and Collaborations, Processes, and Financing. Like any other e-Government models, the eGovSDM also exploits or utilises the three pillars of Principles introduced and discussed by Organisation for Economic Co-operation and Development (OECD) e-Government Project [18] and Brendan [4] as: Pillar I, engage citizens and open up government to maintain public trust. They explained that the actions needed to exploit the potential and minimise the risk of using technology, and new technologies in particular, for open, participatory and ubiquitous public sectors where institutional and non-institutional actors could engage and collaborate with Governments. Pillar II, adopt joined-up approaches to delivering public value. They further discussed that it focuses on the strategic importance of coherent ICT use across the public sector. They added that as the infrastructure of the twenty-first century, ICT systems, standards, and services enable joined-up and agile administrations capable of adopting whole-of-society approaches to create public value. Pillar III, strengthen capabilities to ensure return on ICT investments. They stressed that these were the conditions for ensuring success in ICT decision-making and management.

The eGovSDM would also utilise the above discussed pillars of Principles to enhance its operational elements or strategies in an effort to achieve the Vision 2030.

3.3 Infrastructure/Platforms

In order to attain the Vision 2030, the eGovSDM's approach is to envisage the Principles of the Infrastructure that is a critical e-Governance requirement in this work. The Government has to ensure that state-of-art e-Government Infrastructure is developed and sufficiently deployed across the country including the most difficult areas to be accessed such as remote sites. The NDP3 [25] Report stressed that a well-functioning infrastructure is a pre-requisite for economic activity in virtually all sub-sectors and the attainment of Vision 2030. The Report further emphasised that adequate infrastructure raised productivity, lowered production costs, and helped those in rural and remote areas to participate in the economy.

From the report's emphasis, it is a mandatory and obligation of the Government to ensure that all the relevant Infrastructure is provided in the most disadvantaged communities. Some of the developed Infrastructure include: the wired and wireless networks, data centres, data backup facilities, servers, cloud computing structures, websites, portals or "one-stop-shops", kiosks, network security facilities, virtual private networks (VPN), long term evolution (LTE) technologies, uninterrupted power supply (UPS), 24 h electricity supply, backup electricity generators, solar panels, battery for solar panels, personal computers (PCs), operating systems, laptops, mobile phones, printers, scanners, photocopiers, fax machines, routers, switches, repeaters and many more network devices, communications equipment, but to mention a few. The Government has to consider write manuals that could help the communities from all walks-of-life to use the deployed Infrastructure. The Affordable telecommunications infrastructure and networking should penetrate all spheres of human activity including in homes in both rural and urban areas to transform Namibia into a society where information and knowledge are key assets.

3.4 Capabilities

One of the aspects of Namibia's Road-Map to Vision 2030 was to ensure that the nation adequately trains the citizens in all categories of employment sectors. The country has to heavily invest in training the professionals to acquire high level relevant skills in various technologies such medicine, engineering, networking, research, agriculture, banking, insurance, software development, technicians, mining, aviation, but to mention a few. Much emphasis should be placed on equipping the disadvantaged communities with the basic knowledge to operate the IT/ICT and any other installed systems. According to OECD E-Government Project [18] pointed out that Governments should regularly evaluate the impact ICT emerging technologies, trends and projects had on staff, assess skill gaps and ensure development of new types to enhance organisational learning and match fast changes taking place. The OECD went on to say that, the public sector work force needed to be able to count on flexible skills and competencies. It (OECD) emphasised on the skills for the advanced use of new technologies (social media and mobile technology) in carrying out internal tasks, deliver services and engage with outside actors; those on ICT project management; and those sustaining use of data for policy modelling, evaluation, data analytics and data mining to sustain policy simulation, target improvements in service delivery, monitor public policies' and programmes' impact.

The OECD E-Government Project [18] also stated that in order to ensure availability of adequate skills, governments should leverage the strengths of partners in the private and non-governmental sectors as necessary and be innovative in the ways they tap into high-skilled labour market. It emphasised that arrangements to develop an adequate public sector workforce include recruiting and involving young professionals, creating "centre of excellences," providing professional and

vocational training, establishing exchange programmes and win–win relations between the public sector and technology leaders. The OECD further stressed that strategies for ICT skills should be developed in line with other policies dealing with public sector work force mobility and ageing.

In view of the above, the eGovSDM would drive the Namibia's Road-Map to Vision 2030 by equipping the citizens from all walks-of-life with high level relevant skills across various technologies. This would enable the nation fill-in the employment position with capable workforce. In that way, the workforce would have the capabilities to effectively carry-out duties satisfactory. The eGovSDM would ensure that the Government would across all employment sectors frequently evaluate the impact ICT emerging technologies, trends and projects had on workforce and assess skill gaps. In addition, the Government would ensure that as technology evolves, the workforce also keep on adapting and acquiring new skills to keep pace with trends of current development.

3.5 Access

The eGovSDM's vision is to ensure that the element Access facilitate the Namibia's Road-Map to Vision 2030 by allowing the citizens have rights to acquire and be equipped with e-Government data such as: digital data, re-use public sector data, data interoperability, in-house and external developed content. In that way, the Namibian Government would improve on the business accessibility, transparency, accountability, service design and delivery, decision-making, but to name a few. This was supported by Valentina [27] who pointed out that a vision of e-Government implies providing greater access to information as well as better, more equal services and procedures for public and businesses. He highlighted that even when e-Government projects seek to improve internal government processes, the end goal should be making government serve citizens better. The OEDC (2013) also argued that Governments should adopt legislations that recognised citizens' right to access, use and re-use public sector data, information, records and content. The OEDC (2013) further emphasised that actions should be taken to strengthen a culture of access and use of data to spur participation in policy making, creation of public value, service design and delivery. The OEDC (2013) described the objective of public ICT initiatives as to foster networked societies and collaborative Governments where all actors were given the opportunity to exploit, access and harness services and information anywhere and at any time. Whereas, Rabaiah and Vandijck [19] debated that the third most important guiding principle for e-Government is to achieve universal access. They further emphasised that there was no point in designing state-of-the-art services (online or otherwise) at high costs without being accessible to everyone. They also added that services should be accessible to all, indiscriminately, regardless of their financial abilities, language, geographical location etc.

Other e-Government authors, such as Brendan [4], stated that easy access to public data could also encourage participation and collaboration by citizens and businesses. He went on to discuss that the provision of open data could help to promote accountability and transparency by encouraging new solutions, removing inefficiencies and improving decision-making.

As discussed by other researchers, indeed the eGovSDM would ascertain that the Namibia Government accord its citizens Access to the relevant e-Government data, that they would utilise it to seek solution to twenty-first century developmental problems. In that way the citizens would have capabilities to do their respective work efficiently and able to make concrete decisions.

3.6 Content

Namibia's Road-Map to Vision 2030 is to build the tailor made Content translated into local languages, to give an equal opportunity even to illiterate citizens to read and confidently use it. This is supported by Bruno et al. [5] who pointed out that in many countries, content must be provided in more than one language or dialect. They further argued that e-Government should also address the needs of those who are illiterate. They emphasised that local language and content tailored to be used to different communities. They stressed that content was supposed to be in local languages and that interfaces were easy to use by the illiterate citizen. From these expressions, the eGovSDM's vision is to ensure that the Content is developed and translated into all the Namibian ethnic's languages or dialects to allow every citizen have Access and Capabilities to use it. It this way, even the illiterate citizens from the disadvantaged communities would be able to use such tailored Content.

Bruno et al. [5] again pointed out that published sites, and rich in Content, were just a first step in maintaining e-Government programmes. They described e-Government that it had the potential to involve citizens in the governance process by engaging them in interaction with policymakers throughout the policy cycle and at all levels of government. Strengthening civic engagement contributes to building public trust in government. They further emphasised that the Government should always focus on Content that supported other goals, such as economic development, anti-corruption, attracting foreign direct investment.

Stephen [23] also pointed out that African Governments needed to develop appropriate policy frameworks, supported by legislation for e-Governance, that were linked to strategic development objectives; enlisted high-ranking political e-Government champions; focus awareness, outreach and training efforts on the less privileged segment of targeted users, particularly women and neglected rural communities and promote local Content and supports local language development.

The ITU [13] Report examined what brings users to the site and makes them want to use the e-Government services. The Report further pointed out that such information will help decision-makers develop strategies to increase the use of

an e-Government service by considering ways to make the site easier to use, incorporating more content to meet the needs of the public, marketing and supporting user communities, and addressing issues about ICT access, security, literacy and training. The ITU [13] Report also narrated that typical practices of citizen-centric approach to e-Government included: organising Content around citizen needs; aligning the structure of the pages in the web site to reduce the number of clicks it takes to find information, Access a service, or to complete a transaction; improving the affective qualities of the site; adding functions to facilitate the communication between citizens and the Government; and, enabling the user to customise the site contents. Whereas, Subhajit [24] highlighted that the key to encourage citizens in developing countries to use the Internet would be to provide them with compelling Content and services that met their primary needs. Subhajit [24] also stressed that the solution would be to persuade citizens in developing countries to use the Internet by providing them with convincing Content and services that met their essential needs. He further emphasised that free availability of compelling Content may be the single most important action to boost Internet use and reduce the digital divide.

3.7 e-Applications

For the success of the Namibia's Road-Map to Vision 2030, the eGovSDM should envisage on the heavily utilisation of e-Applications across all walks-of-life. The eGovSDM's emphasis is that all Namibian line-Ministries should actively use e-Applications such as e-Health, e-Education, e-Mining, e-Transport, e-Agriculture, but to mention a few. In that way, these e-Applications would ensure citizen satisfaction in relation to delivery of critical public services like water, electricity, Internet, health care and sanitation. In support of this, the UNESCO [26] Report highlighted the training's aims that sensitised senior and middle-level public servants in line-Ministries and sub-national Governments to the utility of e-Applications in ensuring client satisfaction in regard to delivery of essential public services such as water and sanitation services, health care, education and electricity, etc. The UNESCO Report further narrated additional utilities of e-Applications such as: to ensure integrity and transparency in public procurement in the context of improving the quality of public services to citizens; manage human resources in a coherent and coordinated mode; open, transparent and objective evaluation and selection of bids; enforce standard Governmental financial regulations; carry out needs assessment mapping and build a climate of trust between service providers and consumers. Also, Bwalya and Healy [6] argued that the resource-constrained African countries had also started embracing the concept of digital and knowledge economies which had ushered in an escalation and efficient use of e-Applications such as e-Health, e-Learning, e-Government, etc., in everyday lives of even marginal individuals living in economically excluded places.

3.8 Interaction and Collaborations

The eGovSDM promotes continuous close Interactions and Collaborations in electronic usages among relevant stakeholders from all walks-of-life, which forms a web of relationship. Such an Interactions and Collaborations is perceived to manage electronic transact among the main blocks of the e-Government such as Government to Citizen (G2C), Government to Business (G2B) but to name a few. This is supported by Valentina [27] and Mansell and Wehn [15] who argued that continuous interaction and communication between Government and its stakeholders contributed to the creation of awareness about the potential contribution of ICT to local community activities. They also pointed out that the electronic transactions and interactions between Government and each group constituted the e-Government web of relationships and the respective four main blocks of e-Government, that were: first, Government to Citizens (G2C), second, Government to Business (G2B), third, Government to Government (G2G) and fourth, Government to Employees (G2E). They further emphasised that e-Government played a vital role, not only in facilitating market-led initiatives but also in initiating the process of capability building and in coordinating the actions of a large number of interested stakeholders.

3.9 Processes

The eGovSDM fore-saw that Namibia's Road-Map to Vision 2030 was to collectively coordinate and automate ICTs' Processes for all sectors of work. The eGovSDM emphasis is to manage all technological and social Processes at one-stop as an effective and profitable way to serve the citizens. Carsten [7] pointed out that countries reach more advanced stages of e-Government, and they need to pay attention to the development of coordinated Processes and the allocation of responsibilities. He emphasised that creating a one-stop Government interface was a major challenge in national e-Government efforts. He also stressed that the introduction of e-Government and the integration of services usually require Governments to streamline their administrative Processes. Other scholars, such as Jon [14] argued that e-Government relies on ICTs to automate the Processes to serve citizens, businesses, Governments and other constituents particularly through the Internet and the World Wide Web. He further highlighted that e-Government involves taking computer-based technologies and combining them with human-based administrative Processes to create new ways of serving citizens. He went on to stress that organisations had to adapt ICTs to business Processes, and business Processes had to adapt to ICTs as well.

3.10 Financing

The eGovSDM has to envisage continuous funding schemes with the Government and private sectors in order to sustain the programmes of e-Government. Sam et al. [21] explained that e-Governance programmes spanned a long duration of time, and it was absolutely essential that such programmes were funded continuously to keep them on track. They further argued that many projects that had failed have been victims of non-funding by the Governments. They also pointed out that one prominent alternative to keep such programmes funded for a long time was to have public–private partnership (PPP) that helped private enterprises fund the programme and at the same time enjoyed the results.

3.11 Monitoring and Evaluation

The Monitoring and Evaluation (M&E) element of the eGovSDM serves the purpose to help improve performance of each element and to achieve the results desired. M&E within the eGovSDM will be used to evaluate and measure using a results-oriented approach the outcomes and outputs of each element of the model as well as the overall service delivery performance gains achieved by the different sectors of government. The progress made in achieving objectives for each element and the quality of services provided will be observed on a continuous basis through monitoring. Feedback on progress and performance of service will be provided to relevant political and administrative leadership. Observations made through the monitoring process will be analysed and assessed on a frequent basis to establish the reasons for good or poor performance of each eGovSDM element. The results from M&E will be used to correct areas that underperform.

3.12 Service Delivery Performance

The above discussed eGovSDM's elements or strategies would facilitate the Service Delivery Performance in all walks-of-life. When all the discussed elements or strategies are functional, would enable lengthy procedures cut short, that's reducing bureaucracy, easy access from remote sites, quick transaction, enhance citizen's effective participation, cost reduction and many more. Sangita and Bikash [22] pointed out that the indicators of good governance in service delivery processes were the potential to address efficiency, accountability, participation and equity. They further discussed these in detail as: first the efficiency included speedy delivery, cost reduction and accessibility of service; second, accountability that covered transparency and simplification of procedures and actions, reduction of corruption; third, participation which comprised of empowerment with information

in formulation, implementation and monitoring of policies; and equity, which involved benefits for disadvantaged groups and areas/regions. Other scholars like Rafia [20] argued that the problems of poor service delivery in the developing nations were mostly due to lack of accountability, transparency and commitment in making services work for poor and marginalised citizens. Gordon [10] also cited that in the traditional model of public service delivery in many developing countries, procedures were long and time consuming, and often lack transparency. He gave further examples, such as "a citizen wishing to obtain certificates and records and other official documents similarly must manually complete request forms, make personal visits to multiple offices, and again devote significant time to the process." He also emphasised that e-Government processes, in contrast, offered around the clock access from distant locations, reduced bureaucracy, faster and more convenient transactions, lowered transaction costs for applicants and improved quality of service.

4 Discussion

The elements of the eGovSDM each play an individual role. At the bottom of the eGovSDM model are the Approach and Principles. The Approach and Principles ensured that the Approach in taking plans is well structured and executed to drive the rest of the elements or strategies of the model. For instance, the Approach to prioritise the Finances allocated to the e-Government projects; the type of Infrastructure to use; prioritise the deployment of Infrastructure to the needy or poor citizens; prioritise the campaign awareness to the disadvantaged communities, plan to properly structure the conducive regulatory flame-work; plan for the designing of effective ICT policies, set to cut down bureaucratic systems; plan to build the tailor made Content translated into the local Namibian languages; plan to encourage line-Ministries to use e-Applications across the country; to illustrate a few. Once a comprehensive Approach to planning of undertaking e-Government projects is well done at the initial stage, then the Namibian Government would manage effectively deliver the services to the citizens.

At the bottom, there is the Principle which engages the citizens to participate in the utilisation of any deployed technology across all sectors of work. Such a technological universal engagement will allow the citizen operate any available system and be in position to use the tailor made and translated Content across the systems. In that way, the Principle will make all citizens regardless of the status to have access to the system and the Content.

In the model, second from the bottom there is Monitoring & Evaluation. This is the element where the Namibian Government can take an assessment and inventory of the Approaches and Principles done, Finances used, Infrastructure deployed, tailor made Content used, e-Applications deployed and Processes implanted. This is to ensure that standards are followed to allow the success of rendering services to all citizens of the country.

Also, second from the bottom is the Finance, which is the backbone of the eGovSDM elements or strategies. It is a key and critical in the sense that each and every element or strategy requires adequate funds. Many of the e-Government projects are long term, and to sustain them require continuous funding. For instance, to train the entire country's population on how to use the e-Government systems will need an exorbitant budget. To acquire the state-of-art e-Government infrastructure, ranging from telecommunications to network equipment, also will require huge budget. To develop the tailor made and local translated Content shall need a big budget. In fact, all the elements or strategies of the eGovSDM require huge and continuous funding.

For smooth running of services, each and every line-Ministry will need to design and build e-Applications. Such applications will make line-Ministries be able to transact online. Their systems shall be interlinked and that encourages the interoperability of Content and the hardware.

5 Conclusion

Nations in the sub-Sahara region such as Namibia have set themselves national goals and visions to better and improve service delivery to their citizens. The attainment of these visions and goals is hampered by numerous barriers, such as the lack of infrastructure, funding, trained manpower, resources, poor planning, management, regulatory laws, bureaucratic processes and accountability. These barriers negatively impact on the delivery of services provided to citizens and business by both the public and private sectors.

As a means to address these issues, Namibia has initiated its e-Government initiative and developed a Strategic Plan of Action. The eGovSDM (see Fig. 1) was envisaged to enhance future national development through the utilisation of e-Government benefits, trends and solutions to improve service delivery in the Government and private sectors.

The eGovSDM consists of 11 functional components in this work referred as elements or strategies such as: Approach and Principles, Financing, Monitoring and Evaluation, Infrastructure/Platforms, Access, Capability, Content, Interaction, Collaborations, Processes and e-Applications that work together to improve and extend service delivery.

Through the utilisation of the eGovSDM, Namibia's Vision 2030 the development objective, "Accomplish the transformation of Namibia into a knowledge-based, highly competitive, industrialised and eco-friendly nation, with sustainable economic growth and a high quality of life," can be achieved. The eGovSDM will also facilitate the achievement of Namibia's e-Governance initiative by providing the guidance for improving service delivery.

The generic nature of the eGovSDM makes it suitable for use by any country for the attainment of their National e-Government Strategies.

References

1. Arc.gov, Core Principles of e-Governance (2001), Arc.gov.in/11threp/ARC_11thReport_Ch5.pdf
2. M. Backus, e-Governance in developing countries. IICD Research Brief No 1, March 2001, website: www.iicd.org
3. B.J. Berman, W.J. Tettey, African states, bureaucratic culture and computer fixes. Pub. Adm. Dev. **21**, 1 (2001)
4. H. Brendan, Supporting Public Service Reform e-Government 2012–2015. Minister for Public Expenditure and Reform, 2012
5. L. Bruno, A.K. Mohsen, C. Vivek, The Government Handbook for Developing Countries a Project of InfoDev and the Center for Democracy and Technology, 2002
6. K.J. Bwalya, M. Healy, Harnessing e-Government adoption in the SADC region: a conceptual underpinning. Electronic J. e-Government **8**(1), 23–32 (2010). Available online at www.ejeg.com
7. F. Carsten, Introduction: e-Government Readiness Assessment Framework. International Telecommunication Union E-Government Implementation Toolkit, 2009
8. Economist Intelligence Unit, e-Government in Europe, Middle East and Africa, 2013
9. eGov Namibia Report, Namibia public service e-Government strategic action plan (2013–2018). (Office of the Prime Minister, Republic of Namibia, 2013)
10. W. Gordon, e-Government: Key Concepts and Applications to Development, 2007
11. R. Heeks, Achieving Success/Avoiding Failure in e-Government Projects, IDPM, University of Manchester (2003), Downloaded from http://www.egov4dev.org/topic1smry.htm
12. M.E. Hernán, e-Government Architectures, Technical and Political Situation in Latin America (Items International, ECLAC, Division of Production, Productivity and Management Santiago, Chile, 2007), http://www.eclac.org/SocInfo
13. International Telecommunications Union (ITU) Report, Electronic Government for Developing Countries: ICT Applications and Cyber Security Division Policies and Strategies Department ITU Telecommunication Development sector, 2008
14. P.G. Jon, Electronic Government for Developing Countries. The International Telecommunication Union (ITU), 2008
15. R. Mansell, U. Wehn, *Knowledge Societies: Information Technology for Sustainable Development* (Oxford University Press, Oxford, 1998)
16. S. Odile, Organisation for Economic Co-operation and Development (OECD) e-Government studies future of e-Government AGENDA 2020, in *OECD E-Leaders Conference*, 2000
17. Office of the Prime Minister (OPM), The e-Governance Policy for the Public Service of Namibia (Windhoek, Republic of Namibia, 2005)
18. OECD e-Government Project, Public Governance and Territorial Development Directorate Public Governance Committee. 2013, GOV/PGC/EGOV(2013)1
19. A. Rabaiah, E. Vandijck, A strategic framework of e-Government: generic and best practice. Electronic J. e-Government **7**(3), 241–258 (2009). Available online at www.ejeg.com
20. N. Rafia, E-governance for improved public service delivery in Fiji. J. Serv. Sci. Manag. **3**, 190–203 (2009). doi:10.4236/jssm.2009.23023
21. F. Sam, K. Pradeep, N. Vijaykumar, e-Governance projects: exploring the way to success. SETLabs Brief. **9**(2) (2011)
22. S.N. Sangita, C.D. Bikash, Electronic Governance and Service Delivery in India: Theory and Practice, 2005
23. C. Stephen, African e-Governance – Opportunities and Challenges (2002), Http://www.ansa-africa.net/uploads/documents/.../Coleman_e-Governance.pdf
24. B. Subhajit, e-Government and developing countries: an overview. Int. Rev. Law Comput. Technol. **18**(1), 109–132 (2004). Carfax Publish Taylor and Francic Group
25. Third National Development Plan (NDP3) 2007/2008–2011/12, Republic of Namibia Windhoek

26. United Nation Educational, Scientific and Culture Organisation (UNESCO) Report, African Training and Research Centre in Administration for Development: Curriculum Guide on e-Governance for African Government Institutions (2008)
27. D.N. Valentina, e–Government for developing countries: opportunities and challenges. Electronic J. Inf. Syst. Developing Countries **18**(1), 1–24 (2004). http://www.ejisdc.org
28. E. Zakareya, I. Zahir, e-Government adoption: architecture and barriers. Bus. Process Manag. J. **11**(5), 589–611 (2005)