ISSN 1352-9404

UWS UNIVERSITY OF THE WEST of SCOTLAND

School of Engineering and Computing

Computing and Information Systems Journal

Vol 20, No1, 2016

Edited by Abel Usoro

www.uws.ac.uk

Computing and Information Systems is published normally three times per year, in February, May, and October, by the University of the West of Scotland.

From the next issue the editorial address is Dr Abel Usoro, School of Computing, University of Paisley PA1 2BE; tel (+44) 141 848 3959; fax (+44) 141 848 3542, e-mail: cis@uws.ac.uk or abel.usoro@uws.ac.uk.

Editorial Policy

Computing and Information Systems offers an opportunity for the development of novel approaches, and the reinterpretation and further development of traditional methodologies taking into account the rate of change in computing technology, and its usage and impact in organisations.

Computing and Information Systems welcomes articles and short communications in a range of disciplines:

- Organisational Information Systems
- Computational Intelligence
- E-Business
- Knowledge and Information Management
- Interactive and Strategic Systems
- Engineering
- E-Learning
- Cloud Computing
- Computing Science

The website for Computing and Information Systems is http://cis.uws.ac.uk

Notes for Authors:

Articles for publication should be submitted on disk or via e-mail, in Microsoft Word for Windows format. A style template can be found at http://cis.uws.ac.uk/crow-ci0/articles/style.doc. Please observe the following guidelines:

1. The Normal paragraph style is Times New Roman 11 point, justified, with 6 pt space before. Avoid inserting extra empty paragraphs. Avoid the use of additional spaces between words or after a full stop. Use full stops after initials and precede subsequent initials by a space. Use the standard heading styles Heading 1, Heading 2 etc for top-level and subsidiary headings within the article.

2. Compose figures in an enclosing invisible frame or rectangle, with captions and any text on the figure provided in embedded textboxes, and all including the enclosing rectangle grouped into one object, defined to have "Top and Bottom" text wrapping. Try to avoid using anything based on the Normal paragraph style in embedded text boxes.

3. Equations should be embedded in the text, possibly as paragraphs in their own right. They should not be treated as figures.

4. All entries in the reference section should be cited in the text. List references at the end of the paper, alphabetically by the first letter of the first author's last name. References should be identified in the text of the paper by typing the corresponding name and year in parentheses: letters a, b, etc can be added to the year if several references by the same author for the same year are cited. If a page number is set it should be set as (author, year, page number). DO NOT NUMBER REFERENCES; they must be alphabetical and unnumbered. References should have a half-inch hanging indent as shown below and there should be no additional lines between references. Book titles and names of journals should be printed in italics, not underlined. Examples:

Author, A., Colleague, C., and Friend, F. (1995a). Title of paper, *Journal*, **3** (1): 25-50.

Author, A. and Student, S. (1995b), Another paper, in *Title of Book or Conference*, (E. Editor, ed.), XYZ Press, Place of Publication: 451-457.

Author, A. (1995c). *Title of Book*, XYZ Press, Place of Publication.

Subscriptions:

Computing and Information Systems is issued normally three times per year, in February, May, and October. It is available at an annual subscription of £50, or £20 per issue. Contact the School of Engineering and Computing, University of the West of Scotland, Paisley PA1 2BE. Telephone: (+44) 141 848 3959; Fax (+44) 141 848 3542.

Table of Contents Vol 20, No1, 2016

UWS UNIVERSITY OF THE WEST of SCOTLAND

Printed by Printing Services, University of the West of Scotland
University of the West of Scotland is a registered Scottish charity. Charity number SC002520

An Experimentation of the Use of Oracle PL/SQL with Genetic Algorithms for Bulk Resource Scheduling

Kevin Kelly[i] and Abel Usoro

School of Engineering and Computing
University of the West of Scotland (UWS)
United Kingdom
B00159433@studentmail.uws.ac.uk and abel.usoro@uws.ac.uk

Abstract: This paper investigates the use of Oracle PL/SQL and Genetic Algorithms for performing bulk resource scheduling tasks. To achieve this current theory and information from literature were examined, in order to select a suitable set of methods and algorithms. The findings of this paper are that PL/SQL and Genetic Algorithms when combined with knowledge bases provides an acceptable level of accuracy and speed when scheduling multiple of constrained resources.

Keywords: *Oracle PL/SQL, Nearest-Neighbour, Genetic Algorithm, Knowledge Database, Recourse Scheduling, Appointments*

1.0 Introduction

Computerised scheduling applications are used within a variety of different businesses. These applications help to improve schedule accuracy and free office resources from time intensive tasks. One such application that the author has industry experience performs a scheduling routine on a night basis. The general performance of this routine suffers during peaks in the business during the winter months. Given that the inputs and outputs of this routine are known the author began to investigate possible methods within the companies infrastructure that could over a routine with better performance. This paper therefore seeks to answer the question: *Can a scheduling system based on Oracle utilizing PL/SQL outperform that of a 'black box' solution that is already in use in terms of bulk resource scheduling?*

Through a detailed review of the current theory published by Oracle and methods and algorithms published in current literature a PL/SQL application was developed and comparable tests were conducted for this paper. The rest of the paper will present Oracle PL/SQL, genetic algorithm, knowledge base, PL/SQL pseudo code used, results, discussions, and conclusions.

2.0 Oracle PL/SQL

Oracle Procedural Language extensions to SQL (PL/SQL) is a database technology that was developed by Oracle for use with their various database products and can be found in many of their other products and applications. PL/SQL contains key programming features such as simple IF, ELSE conditions, variables, loops, functions and procedures. PL/SQL also supports more advanced programming concepts such as different types of arrays, advanced loops as well as many concepts found in object-oriented programming languages.

(Oracle, 2015) states that the PL/SQL engine accepts procedural statements as well as standard SQL statements in the form of PL/SQL blocks. Statements are then processed either by the Procedural Statement Executor or the SQL Statements Executor.

PL/SQL programmes are often compiled and stored within the database itself; this is a key feature of this programming language. The purpose of this is to lower Input/output (I/O) calls and increase portability with other Oracle products. Although it is possible to store PL/SQL within the database it is also possible to run a PL/SQL script or application against the database without storing it.

The application in this paper is stored within a database package. Packages allow for a group of statements to be contained within a single entity.

3.0 Genetic Algorithm

Genetic Algorithms (GA) are a form of meta-heuristic algorithms. (Brezulianu, et al, 2009) demonstrates the use of GA to solve a constraint based scheduling problem for employees working within a retail environment which includes multiple different shift patterns. By placing each of the resources and their individual constraints into arrays (Chromosomes) it allows

the information to be quickly accessed and stored in a single location. (Brezulianu, et al, 2009)further details the construction of a custom chromosome for use in their project needs, containing the following values:

- Shift start time
- Shift duration
- First free day
- Second free day
- Employee Department.

The ability to customize chromosomes towards a specific problem is very advantageous.

3.1 Generations

Generations refers to the number of iterations the genetic algorithm will run or loop for. New populations are created for each iteration and then discarded, only retaining the 'elitist' chromosomes which are stored in a constant population. Discovering the correct or optimum number of iterations used in the ‘generations' loop is a time consuming practise; it relies heavily on the size of the population. The number of generations should either be increased or decreased depending on the population size. (Gotshall et al) details various experiments with an increasing and decreasing population size and a varying generation size. This researcher states that it is difficult to find an optimum size for both the generation and population variables. Given more generations, there is a greater chance of a probability events occurring such as crossover or mutation. The number of generations is not always a fixed value; multiple sources in literature including (Raja et al, 2013) and (Gotshall et al) discuss the exploration and the benefits of a fixed generation size versus a dynamic generation size.

3.2 Chromosome

A chromosome represents a grouping of values held within a single structure much like an array. Each chromosome could be considered as a potential solution to a problem statement. As previously detailed by (Brezulianu et al, 2009) the structure can be customized to suit the needs of a specific problem. Given this, the designing of a chromosome is an extremely important process when designing a genetic algorithm (GA). The process of populating these structures is known as encoding. The structures of the chromosomes used in this paper are detailed below:-

- CHROMO_ID – Unique identifier used as the primary key allowing the database to quickly read the chromosome from memory.
- TOTAL_DIST – Individual chromosomes total distance based on appointments.
- FITNESS - Individual chromosomes fitness.
- APPT_ID_1 – Appointment slot 1.
- APPT_ID_2 – Appointment slot 2.
- APPT_ID_3 – Appointment slot 3.
- APPT_ID_4 – Appointment slot 4.
- APPT_ID_5 – Appointment slot 5.
- APPT_ID_6 – Appointment slot 6.
- APPT_ID_7 – Appointment slot 7.
- APPT_ID_8 – Appointment slot 8.

3.3 Encoding

Encoding is the name given to the process of creating the individual chromosomes. The type of encoding that this paper makes use of is known as 'permutation encoding'. It is reported (Obitco, 1998) that this form of encoding is quicker for scheduling problems than non-permutation encoding. This type of encoding lowers the overall search space of the GA. This statement has evidence in current literature and is discussed by (Anderson, 2011), that it increases overall throughput significantly.

The problem that this paper addresses involves appointments which could either be AM (in the morning) or PM (afternoon and evening). An example of a chromosome structure is shown in the Table 1.

Table 1 Basic Chromosome Design

AM				**PM**			
Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Slot 6	Slot 7	Slot 8
8 - 9 am	9 - 10 am	10 - 11 am	11 - 12 am	12 - 13 pm	13 - 14 pm	15 - 16 pm	16 - 17 pm

In order to further optimise the encoding process the PL/SQL application treats the AM and PM separately by splitting the chromosome into two separate sections. These are shown in Table 1:

appointments slots 1 to 4 are treated as AM appointments only and appointment slots 5 to 8 are treated as PM appointment only. This allows the chromosomes to handle the problem constraints of appointment slots for different periods of the day.

The research further optimises this process by using a simple nearest-neighbour search heuristic that attempts to weigh appointments based on their distance from one another. The purpose of this process is to lower the search space as much as possible. This is known as a greedy algorithm, meaning that they will use the majority of the best solutions on a first come first served basis. A later grouping will have a poorer choice of appointments to attempt to schedule. (Johnson, 2008) states that nearest-neighbour is typically found in a well-defined search space. This paper uses a data set of pre-processed distance information held in a database table. By structuring this information this manner all known answers are readily available and quickly accessible.

There are negative aspects from the use of a nearest-neighbour algorithm. Previously reference was made that it could be considered a greedy algorithm. A further negative aspect documented by (Roeva et al, 2013)is that there is a requirement for creating several instances in order to find the true optimum result. By making use of a pre-defined data set such as a knowledge base that contains all possible results, this problem is removed; there is no need for multiple instances.

3.4 Population

Like generations, population size is an important factor to any GA. If the population size is too low then non-desirable results may be returned; and if the population size is too large then the time taken to produce results can greatly increase.

Like generations, it is important to find a good balance. (Roeva et al, 2013) explains that given a large enough initial population; doubling this population may not increase the chances of producing better results. Through a series of experiments, (Roeva et al, 2013) gives the detail that finding an optimum population size is not a trivial task.

(Ma et al, 2008) investigates different methods of producing an optimum populating size. One of the recommendations from this paper is a method known as 'Random Fluctuation Population'. This method either increases or decreases the size of the population depending on the problem size; this paper will make use of this method. (Ma et al, 2008) concludes that fixed sized populations are wasteful as the optimum may be found at an early stage of the GA.

3.5 Natural Selection

After the creation of the population the first of a group of different operators is performed against this initial population. The first of these is referred to as '*Natural Selection*' which involves selecting the 'fittest' chromosomes from the population. Individual fitness values are calculated for each chromosome.

A selection method is then applied to the population as it stands, making use of the previously calculated fitness value; the purpose of the selection method is to select chromosomes with an acceptable level of fitness and add them to a new population selected by fitness. (Koenig, 2002) explores the choice of selection methods; this paper makes use of 'Roulette Wheel Selection' which is described below.

3.6 Roulette Wheel Selection

Roulette wheel selection is based on probability. A chromosome is selected a random based on a randomly generated value between 0 and 1; any chromosome that fall in this group may be selected at random. This process is repeated until the number of iterations is equal to the total population.

At the end of this process a further population has been created of the fittest chromosomes from the initial population.

3.8 Crossover

A further operator is then applied to this newly created population, this is known as 'Crossover'. Crossover operators are triggered based on a probability calculation; they do not occur on every iteration of the GA. This paper makes use of a method known as 'Single Point Crossover' and is explained below.

3.9 Single Point Crossover

Single Point Crossover works by selecting a random point in a randomly selected chromosome that is then referred to as the parent 1.

As illustrated in Table 2, the appointments held in slots 5 to slot 7 are selected.

Table 2 Crossover Parent 1

Not Selected				**Selected**			
Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Slot 6	Slot 7	Slot 8
1001	1002	1003	--	1005	1006	1007	--

A second chromosome is then randomly selected from the same population, although in this case the other half of the chromosome is selected for crossover. This is illustrated below in *Crossover Parent 2* (Table 3):

Table 3 Crossover Parent 2

Selected				**Not Selected**			
Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Slot 6	Slot 7	Slot 8
1001	--	1003	1004	1005	1010	1008	--

From this selection two possible child chromosomes will be created. These are illustrated below in Tables 4 and 5.

Table 4 Child 1

Parent 2				**Parent 1**			
Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Slot 6	Slot 7	Slot 8
1001	--	1003	1004	1005	1006	1007	--

Table 5 Child 2

Parent 1				**Parent 2**			
Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Slot 6	Slot 7	Slot 8
1001	1002	1003	--	1005	1010	1008	--

After the creation of both child chromosomes, the total distance of each is calculated. If these distances are less than that of their parent chromosomes then the newly created child chromosomes are then written to the selection population, otherwise they are discarded.

3.10 Mutation

Mutation is another type of operator. Again this is triggered on the basis of probability calculation. Mutation adds randomness to the selected population. By adding a form of randomness it allows the GA to explore results that had not occurred during the initial or selection population creation, or which are unlikely to occur during crossover.

This paper makes use of 'Single Point' mutation. A single point is randomly selected within a given chromosome and switched with its neighbour to the right. The mutation operator handles the time constraints of AM or PM appointments by treating the two halves of the chromosome as separate areas. This process is shown in *before* and *post* mutation diagrams (Tables 6 and 7)

The total distance cost of this mutated chromosome is again compared with the original chromosome. If the mutation has produced a chromosome with a lower overall cost it is written to the 'natural selection population'; otherwise it is discarded.

Table 6 before Mutation

Mutation			**Mutation**
Slot 5	Slot 6	Slot 7	Slot 8
1005	1006	1007	1009

Table 7 post Mutation

Mutation			**Mutation**
Slot 5	Slot 6	Slot 7	Slot 8
1009	1006	1007	1005

There are many approached to mutation operators, and interesting example that was not explored in this paper was put forward by (Libelli and Alba, 2000). This author recommends the use of an 'adaptive mutation' based on the fitness of the individual chromosome rather than a probability calculation. The method detailed performs multiple mutations to a single chromosome to lower the requirement for a larger population.

3.11 Elitism

The elitism operator is the final stage within the generation iteration; it occurs for each and every generation. This operator selects the best or fittest chromosomes and stores them in another new 'elite population'. Once all iterations are complete the best result from this population could be considered the closest to optimum solution.

4.0 Knowledge Base

This paper details the use of a knowledge base or a store of information. A knowledge base can be defined simply as a store of structured and unstructured information. This concept is detailed in other papers such as New Operators of Genetic Algorithms for Travelling Salesman Problem (Ray et al, 2004), Section of Good Practices for Small Software Development Teams: A Knowledge Based Approach (Castro et al, 2013) and A Knowledge-Based Multiple - Sequence Alignment Algorithm (Nguyen and Pan, 2013). The application of simple pre-defined rules or data gathering techniques allow for data to be modelled and stored in a meaningful structure; this structure is a knowledge based.

Such a structure was developed for this paper. It contained details of pre-calculated distances between a starting location and each possible end destination as shown in Table 8 as *An Example of a Knowledge Base*.

By building up such a store of data the application was able to utilize this information using simple select query statements, and removing any need for a further calculation.

Table 8 An Example of a Knowledge Base

Starting Post Code	**Destination Post Code**	**Distance**
G1 1ZE	G1 1ZE	0.00
G1 1ZE	G1 5DB	2.12
G1 1ZE	G1 4RJ	1.84
G1 5DB	G1 1ZE	2.12
G1 5DB	G1 5DB	0
…	...	…

To gather this data a suitable distance formula was required and the chosen method is known as the Harversine formula. The Harversine formula is a formula that allows for distances between two locations to be calculated, providing latitude and longitude values; both locations are known. The Harversine formula can be found in published literature including Info Sphere Streams for Scalable, Real-time, Transportation Services by (Bienn et al, 2010). These detail the use of the Harversine formula to calculate all possible distances in the Stockholm road network.

It is noted that the formula is resource intensive, as this project will make use of a central store of information. These calculations will only need to be conducted once with the results stored in a

simple indexed structure. Due to this and the ease of storing information within a database application, the negative performance implications found by (Bienn et al, 2010) can be overlooked for this paper.

The Harversine formula is defined as an equation that is used to calculate the distance between two points using longitude and latitude values. (Igis Map) details the Harversine formula as:

$$a = \sin^2(\Delta latDifference/2) + \cos(lat1).\cos(lt2).\sin^2(\Delta lonDifference/2)$$

$$c = 2.\text{atan2}(\sqrt{a}, \sqrt{(1-a)})$$

$$d = R.c$$

5.0 PL/SQL Pseudo code used

The pseudo code for the PL/SQL GA used in this paper is detailed below:-

- Per Resource loop
- Create initial population with Nearest-Neighbour Heuristic from Knowledge base
- Encoding of problem specific chromosomes
- While generation count < I do Loop
  - Roulette wheel selection
  - Single point cross overcome
  - Swap Mutation
  - Elitism
- End Loop
- End Loop

Each of the operations used in this GA has been previously discussed throughout this paper. This GA was created for the sole purpose of answering the given research question: *Can a scheduling system based on Oracle utilizing PL/SQL outperform that of a 'black box' solution that is already in use in terms of bulk resource scheduling?*

6. Results

The following experiments were conducted using the PL/SQL against known results created from the industry standard scheduling application. All data used in these experiments were created from the application and had no manual interaction.

Experiment 1: Measure time taken scheduling appointments from the front end of the standard scheduling system.

This experiment attempts to identify the delay between the various comportments that make up the scheduling system. This data was also used against the PL/SQL GA.

Table 9 Experiment 1 Summary

	Standard	**PL/SQL GA**
Mean	7.9	0.4
Standard Deviation	9.4	0.0
Variance	119%	1%

Results from *Experiment 1 summary* (Table 9) show that industry standard scheduling system takes on average takes 7.9 seconds to schedule a appointment. Conversely, the PL/SQL GA took 0.4 seconds to perform this task. The standard deviation of times in the standard application is 9.4 seconds whereas the deviation in the PL/SQL GA is 0. The variance of results from standard shows a variance of 119%, whereas in the PL/SQL GA variance is 1%.

Experiment 2: Compare time taken scheduling a known schedule from standard scheduling system.

This experiment compared the speed of the PL/SQL GA against completed scheduling runs of the standard scheduling system. Measures of time in the standard scheduling system are in whole seconds, whereas the PL/SQL application shows time in tenths of seconds.

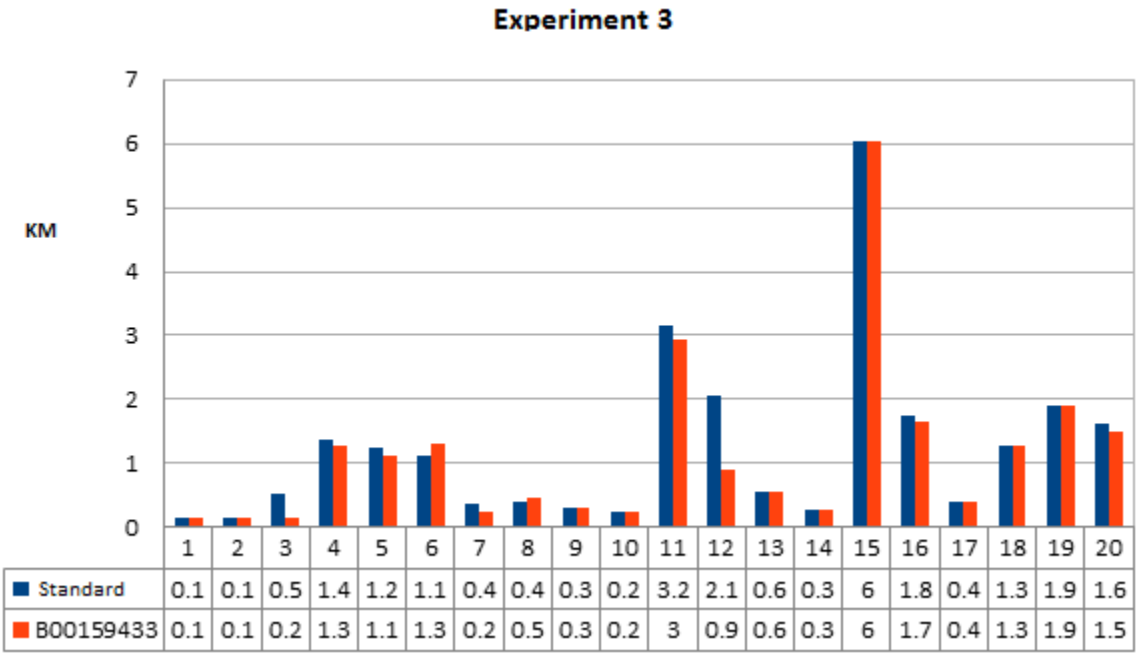
Table 10 Experiment 2 Summary

	Standard	**PL/SQL GA**
Mean	1788.7	11.3
Standard Deviation	1190.9	2.50
Variance	67%	22%

Experiment summary 2 (Table 10) shows that industry standard scheduling system on average takes 1788.7 seconds to produce a schedule. Whereas PL/SQL GA takes on average 11.3 seconds, 158 times slower with comparable data sets. The standard deviation of the standard application is 1190.9 seconds whereas the PL/SQL GA is 2.50 seconds. The variance in the results from standard application 67% and variance from PL/SQL GA is 22%.

Experiment 3: Compare routing accuracy

against known routes prepared by the standard scheduling system.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Standard	0.1	0.1	0.5	1.4	1.2	1.1	0.4	0.4	0.3	0.2	3.2	2.1	0.6	0.3	6	1.8	0.4	1.3	1.9	1.6
B00159433	0.1	0.1	0.2	1.3	1.1	1.3	0.2	0.5	0.3	0.2	3	0.9	0.6	0.3	6	1.7	0.4	1.3	1.9	1.5

Figure 1: Total Distance Chart

The *Total Distance Chart* (Figure 1) displays the total distance of 20 test routes. Routes contain 2 or more sets of appointments; of the 20 data sets tested the following results have been concluded:

- Standard Application - 2
- PL/SQL GA - 8
- Same Result – 10

Both methods created the same results for exactly half of the given test set. The PL/SQL GA created 8 results that were better than that of standard application, and finally the standard application created 2 results that improved that of the GA.

7. Discussions

The results from the first experiment clearly show that there are noticeable delays between the various components that make up the standard application; the PL/SQL GA is contained in a single layer within the database. This gives the PL/SQL application a clear benefit on overall speed; this advantage on overall speed is not only shown in the first experiment it is also clearly visible in the following experiment.

The second experiment only compares time taken by either application to produce an acceptable result. It is not known how the industry standard application calculates its schedules but it is apparent that this application is vastly slower than the one created by this research using Oracle and GA. It could be stated that the standard application is conducting an exhaustive search. (Thomas, et al) states that this is a fruitless search as it is on average 158 times slower than a comparable application using the same data set and performing the same task. If time taken is not a factor then the accuracy by either application is a deciding factor.

The third set of experiments show that on half of all tests conducted both applications obtained the same result, whereas the PL/SQL got a better result 8 of the remaining 10 tests. This could be explained that the applications are using a different method to calculate distances as the differences are small. However the method used in this paper is very accurate for calculating as the crow flies.

8. Conclusions

In conclusion this paper has demonstrated that a database technology when combined with modern algorithms and methods can outperform an industry standard application. Further research and design should be given to the nearest-neighbour algorithm to limit its greedy nature.

References

Anderson M (2011) Industrial Scheduling with Evolutionary Using a Hybrid Representation Available from: Link https://www.diva-portal.org/smash/get/diva2:456830/FULLTEXT01.pdf Accessed: 29th October 2015]

Bienn A, Bouillet E, Feng H, Ranganathan, Riabov A, Verscheure O (June (2010) IBM InfoSphere Streams for Scalable, Real-Time, Intelligent Transportation Services [Online] Available from: http://delivery.acm.org/10.1145/1810000/1807291/p1093-biem.pdf?ip=94.2.219.205&id=1807291&acc=ACTIVE%20SERVICE&key=C2D842D97AC95F7A.86B481D97A46A079.4D4702B0C3E38B35.4D4702B0C3E38B35&CFID=541632871&CFTOKEN=43697059&__acm__=1440925337_b9a1397aa61363c7edd0aaa66bbf6cf9 [Accessed: 30th August 2015].

Brezulianu A, Fira M, Lucian F (2009) Agentic algorithm approach for constraint employee scheduling problem as applied to employees at mall type shops [Online] ACM p.497 - 498 Available from: http://dl.acm.org/citation.cfm?id=1645085&dl=ACM&coll=DL&CFID=543836259&CFTOKEN=17643033 [Accessed: 17th August 2015]

Castro M D Ronney, Braga L Jose, Soares Santos L (2013) Section of Good Practices for Small Software Development Teams: A Knowledge-Based Approach Available from: http://dl.acm.org/citation.cfm?id=2532804 [Accessed: 30th August 2015]

Gotshall S, Rylander B. Optimal Population Size and the Genetic Algorithm. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.2431&rep=rep1&type=pdf [Accessed 9th November 2015]

IGIS Map Haversine formula – Calculate Geographic Distance on Earth http://www.igismap.com/haversine-formula-calculate-geographic-distance-earth/ [Accessed 16th November 2015]

Obitko. Marek (1998) Introduction to Genetic Algorithms http://www.obitko.com/tutorials/genetic-algorithms/encoding.php [Accessed: 29th October 2015]

Johnson C G (April 2008) A Design Framework for Meta-heuristics Available from: https://kar.kent.ac.uk/14451/1/Design_Principles_for_Metaheuristics.pdf [Accessed: 29th October 2015]

Koenig A (November 2002) A Study of Mutation Methods for Evolutionary Algorithms. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.2465&rep=rep1&type=pdf [Accessed 9th November 2015]

Libelli S and Alba P (2000) Adaptive Mutation in Genetic Algorithms Available from: http://www.dsi.unifi.it/~marsili/Papers/Mutation_GA.pdf [Accessed 9th November 2015]

Ma Z, Kring A (2008) Dynamic Population in Genetic Algorithms Available from: http://www2.cs.uidaho.edu/~krings/publications/ACM-SAC-EC-2008.pdf [Accessed: 8th November 2015]

Nguyen D Ken, Pan Y (2013) A Knowledge-Based Multiple-Sequence Alignment Algorithm Available from: http://dl.acm.org/citation.cfm?id=2564682 [Accessed: 30th August 2015]

Oracle (2015) PLSQL Diagram Available from: http://docs.oracle.com/cd/B19306_01/appdev.102/b14261/overview.htm#i4603 [Accessed: 6th September 2015]

Raja P, Bhaskaran M (August 2013) Improving the Performance of Genetic Algorithm by Reducing the Population size. Available from: http://www.ijetae.com/files/Volume3Issue8/IJETAE_0813_13.pdf [Accessed 9th November 2015]

Ray S S, Bandyopadhyay S, Pal K S (2004) New Operators of Genetic Algorithms for Traveling Salesman Problem. Available from: http://shubhrasankar.tripod.com/cgi-bin/tsp1.pdf [Accessed: 6th September 2015]

Roeva O, Fidanovoa S, Paprzycki M (2013) Influence of the Population Size on The Genetic Algorithm Perofmrance In Case of Cultivation Process Modeling. Available from: http://annals-csis.org/proceedings/2013/pliks/167.pdf [Accessed: 8th November 2015]

[i] Kevin Kelly did his MSc Project under the supervision of Dr Abel Usoro. This paper is part of his work.

A Fuzzy Logic Expert System for the Prediction of Breast Cancer

Gabriel Opeyemi Ogunleye
Department of Computer Science
Redeemer's University
Ede, Osun State, Nigeria
ogunleyeg@run.edu.ng

Stephen Gbenga Fashoto
Department of Computer Science
Kampala International University
Kampala, Uganda
Stephen.gbenga@kiu.aac.ug

Emmanuel Weje
Department of Computer Science Redeemer's University
Ede, Osun State, Nigeria
wejee@run.edu.ng

Abstract: Cancer is a serious disease that occurs as a result of genetic mutation in the body. It happens due to the formation of mass tissue known as tumor in the body. This paper develops a fuzzy logic prediction system which serves as a platform for detection of any occurrence of breast cancer in women. The patient's medical history and some risk factors are collected and the results are displayed to show any trace of breast cancer in women. The predictions are made based on the risk factors' inputs which are used to determine the patient's risk.

Keywords: *Cancer, Breast Cancer, Risk Factors, Expert System, Fuzzy Logic*

1.0 INTRODUCTION

Cancer is a dangerous disease which is inherently caused by environmental factors that mutate genes encoding critical cell-regulatory proteins (Malcom, 2001). The pathogens that cause cancer are also known as carcinogens which can be seen in the air, food, water, chemicals and also sunlight that people are exposed to. The World Health Organization (WHO) recorded that 24.6 million people live with cancer.

Breast cancer is a disease that can be easily or commonly found in women today in the world. Numerous women undergo different tests if they have any trace of cancer. These tests are mostly used to ascertain if these women have the Breast Cancer 1 (BRCA1) and Breast Cancer 2 (BRCA2) cancer genes. BRCA1 and BRCA2 are human genes that produce tumor suppressor proteins. These proteins help restore damaged DNA and therefore, play a vital role in ensuring the stability of the cell's genetic material. When either of these genes is mutated, such protein product is either not made or does not function correctly.

However, since mutations in these genes account for only a little fraction of all breast cancers, a negative test result does not mean the women who go for these tests are free from cancer. Likewise, testing positive does not guarantee that the woman will have a breast cancer in the future.

According to Fatimah (2009), 650,000 people out of the estimated 965 million are diagnosed with cancer annually and the lifetime risk of an African woman dying from cancer is two times higher than the risk of a woman in a developed country.

NIGERIA	Male	Female	Both sexes
Population	84,398,000	82,231,000	166,629,000
Number of new cancer cases	37,400	64,700	102,100
Age-standardised rate (W)	79.0	121.7	100.1
Number of cancer deaths	30,900	40,600	71,600
Age-standardised rate (W)	67.4	78.0	72.1
5-year prevalent cases, adult population	67,000	165,000	232,000
Proportion (per 100,000)	139.8	348.6	243.6
First 5 most common cancers			
	Prostate	Breast	Breast
	Liver	Cervix uteri	Cervix uteri
	Non-Hodgkin lymphoma	Liver	Liver
	Colorectum	Colorectum	Prostate
	Pancreas	Non-Hodgkin lymphoma	Colorectum

Figure 1: An estimate of cancer incidence, (adapted from Globocan, 2012)

Figure 1 shows the cases of cancer incidents in Nigeria ranging from number of new cancer cases to the five prominent ones that are common in both sexes in this African region.

Cancer is not only disease that exists; it is just a common disease out of a set of more than 200 different diseases that exist. It can be in a widely accepted sense described as an uncontrolled

growth and spread of abnormal cells in the body. A cell is the basic or fundamental unit of life. Cells in living thing breaks-up to yield more cells when the living thing accommodating them requires it. At some points, cells keep on splitting or dividing without control thus creating excess cells which are not required by the body. This in turn causes the formation of mass tissue also known as a tumor. Tumors can be discovered in tissues of different types. These tumors can either be benign or malignant.

Benign tumor is one in which the tumor does not penetrate its surrounding tissue or spread round the body while malignant tumor is one that may penetrate its surrounding tissue or spread round the body (Artem, 2006). Benign tumors are not regarded as a cancer. They have a slower growth rate than malignant tumors and are usually removed and do not re-occur. Benign tumors are not usually a threat to life. Examples of benign tumors are moles (nevi) and uterine fibroids.

Malignant tumors are cancerous and can be a dangerous threat to life. They invade and damage organs and tissues near them. Malignant tumors develop rapidly, without control and independently from the tissue it emanated from. Examples of malignant tumors include breast cancer, lung cancer and prostate cancer.

Cancer can take place among the people of all ages. It is mostly common in people of over 60 years old and records have shown that one of every people will develop cancer at a particular time in their lives. Cancer has a couple of symptoms. Some of them are listed below:

- Unexplainable weight loss
- Strange bleeding or discharge
- Tenacious indigestion

Detecting the occurrence or traces of cancer early enough can go a very long way in increasing the chances of survival. Some of the techniques for detecting cancer are listed below:

- Self-examinations
- Biopsy
- Ultrasound
- Magnetic Resonance Imaging (MRI)
- Computed Tomography (CT)

There are many proven reasons that make cancer prevail. These reasons can vary from increased ability to take care of diseases which will cause a delay in their progression to inability to diagnose and treat a disease which leads to death and disability (Crimmins, et al., 1994). Predicting the future occurrence of cancer is a very much wanted technique which will be a very great help in providing early detection of cancer and in like manner increase the rate of survivability.

This paper aims at using fuzzy logic to predict the risk of breast cancer occurring in humans based on certain factors therefore increasing the early detecting and the rate of survivability.

The paper is implemented using the Mamdani fuzzy logic approach in MATLAB. A Graphical User Interface (GUI) is also developed using Microsoft Visual Studio 2012. The GUI is powered by a Fuzzy Logic Library and the application uses a Visual C# as its programming language.

The rest of the paper is organized as follows: related works are presented in section 2; section 3 presents the methodology; section 4 presents results and discussions; and conclusions and recommendations are drawn in section 5.

2.0 RELATED WORKS

Research has shown that breast cancer is the most common cancer among women; excluding non melanoma skin cancers. This cancer affects one in eight women during their lives. It occurs in both men and women, although male breast cancer is rare and not common. Breast cancer is a malignant tumor that has grown from cells of the breast. Although scientists know some of the risk factors (i.e. ageing, genetic risk factors, family history, menstrual periods, not having children, obesity) that increase a woman's chance of developing breast cancer, they are yet to find out the actual cause of breast cancer or exactly how some of these risk factors cause cells to become cancerous. Research is ongoing to learn more and scientists are making numerous efforts in understanding how certain changes in DNA can cause normal breast cells to become cancerous.

Early detection of breast cancer improves the chances of a curative treatment and a lot of progress has been made in this field during the last decade. Screening and diagnosis techniques have enhanced early detection of breast cancer over the years. Some of these screening and diagnostics techniques include:

Breast self-examination: The American Cancer Society (ACS) stated that every woman should start monthly breast self-examination at the age of 20. An analysis carried out on 12 studies involving 8,118 patients with breast cancer correlated the performance of breast self-examination with tumor size and regional lymph node status. Women who performed breast self-examination were more likely to have smaller tumors and less likely to have axillary node metastases than those who did not. This technique has been flawed due to the fact that it is rarely performed well.

Clinical breast examination: Women are urged to begin clinical breast examination at age 20. This was stated by the American Cancer Society. They are also advised to have an examination every 3 years between ages 20 and 39.

Mammography: The American Cancer Society, the American College of Radiology, and the American Medical Association each has since updated their guidelines since 1997 and recommended annual mammography beginning at age 40. The National Cancer Institute (NCI) also updated their guidelines in 1997. They recommended that women should undergo screening mammography every 1-2 years beginning in their 40s.

There are many models that have been developed and used to predict the risk of breast cancer in women in the world today. The Gail model is one of such models.

Gail's model was developed by Dr. Mitchell Gail, a senior investigator in the Biostatistics Branch of NCI's Division of Cancer Epidemiology and Genetics (Gail et al.,n.d). Gail's model uses a woman's own personal medical history, her own reproductive history, and the history of breast cancer among her first degree relatives (mother, sisters and daughters) to estimate her risk of developing invasive breast cancer over specific periods of time. The Gail model has been mainly used to test large populations of white women and has been shown to provide accurate estimates of breast cancer risk. The model has also been tested with data gotten from the Women's Health Initiative for African American women. Besides the outstanding performance of the Gail model, its functionality has been limited to white women, and was recently been updated to predict risks for African-American and Asian-American women. The Gail model is not really functional for other racial/ethnic groups apart from the ones given above.

A system that contains three basic modules has been proposed namely; the diagnostic module, the staging module and the treatment recommendation module (Marco et al., 2014). In order to detect patient disease, the patient inputs his/her signs and symptoms to the diagnostic module, which detects what type of cancer the patient has. Once the type of the cancer is established, the staging module finds the current stage of the cancer based on the cancer type and the signs and symptoms provided by the patient. Based on the determined cancer type and cancer stage, the treatment recommendation module can recommend a specific treatment for the case at hand.

Four types of filtering for pre-processing had been developed which mainly concentrate on the means square error (MSE), peak signal to noise ratio (PSNR), Structural (SC) and normalized absolute error (NAE) (Ramani et al., 2013). The authors compared the simulated output parameters such as image quality, mean square error, peak signal to noise ratio, structural content and normalized absolute error. The comparison of four types of filters is tested for 322 mammogram images (MIAS), from the output observation. They concluded that the adaptive median filter is a more appropriate method for detecting breast cancer while compared with other filters, because image quality of adaptive median filter is better than others.

2.1 Fuzzy Logic

Fuzzy logic was introduced by Zadeh (1965), a professor of Computer Science at the University of California in Berkeley. It is aimed at a formalization of modes of reasoning which are approximate rather than exact (Zadeh, 2004). Essentially, fuzzy logic is a multivalued logic that allows intermediate values to be defined between conventional evaluations like true/false and high/low. (Hellman, 2001). Fuzzy logic attempts to systematically and mathematically emulate human reasoning and decision making. It represents an excellent methodology to draw acloseness with computational logic and the human mode of reasoning (Navjot, et al., 2013).

Zadeh (1994) stated that fuzzy logic is a logical system that aims at a formalization of approximate reasoning which is rooted in multivalued logic.

2.2 Fuzzy Logic as a Prediction Tool

Prediction according to the Encarta dictionary is defined as the statement about the future i.e. a statement of what somebody thinks will happen in the future. Predicting the occurrence of an event requires imprecise and uncertain knowledge. Predicting a system is normally done by getting knowledge from past events for which historical data is gotten and analyzed to get to know the resulting sequence or pattern (Justin and Eric, 2006). Knowledge from the past is applied mainly to acquire the patterns that formerly existed. Getting knowledge about the past can provide to some extent knowledge about the future.

Fuzzy logic has emerged over_time as a valuable technique in predicting the occurrence of future events (Vaidehi et al., 2013). It has been used and still is being used as a prediction tool for many systems since its inception in 1965. Fuzzy logic is almost perfect for the prediction of an event because prediction requires knowledge that is rather imperfect and uncertain. The accuracy of a fuzzy logic prediction system is highly dependent on the variables which include the system, expert rules and the membership functions (Michael, 2010). A fuzzy model operates on the level of linguistic terms and it also represents and processes uncertainty. There are many areas in which fuzzy logic has been used; some of these areas are discussed in the next section.

2.1.1 Benefits of Fuzzy Logic as a Prediction System

According to Yilmaz and Ayan (2011), fuzzy logic has the following benefits as a prediction system:

- Flexible because it helps in representing knowledge in an environment of uncertainty and imprecision
- Tolerant of imprecise data
- Suitable to determine bottlenecks within a system
- Models nonlinear functions of arbitrary complexity
- Can be built on top of the experience of experts
- Provides real-time results
- Conceptually easy to understand and is a more intuitive approach
- Based on natural language i.e. built on the structures of qualitative description used in everyday language therefore making it easy to apply

2.1.2 Applications of Fuzzy Logic for Prediction

Water Consumption Prediction of Istanbul City

Some authors conducted a research using the Takagi Sugeno fuzzy method to predict future water consumption of Istanbul city in Turkey using three antecedent water consumption amounts (Abdusselam et al., 2005). The fuzzy model used in this research was made up of three input variables and one output variable. The input variables consisted of the three water consumption values and the output variable indicated the current water consumption variable.

Data was gotten from the Sugeno fuzzy time series analysis which was applied to the monthly water consumption data of Istanbul city. There was a nine years' period of uninterrupted monthly water consumption records. This nine years period ranged from 1995 to 2004. The record period was separated into two periods namely, training periods and prediction periods. The training periods ranged from 1995-2002 while the prediction periods lasted for 18 months.

Development of a Fuzzy Logic Based Rainfall Prediction Model

A research work was carried out on the ability of fuzzy logic in modeling rainfall for South Western Nigerian. The fuzzy logic model that was developed comprised of two functional components which are the fuzzy reasoning and the knowledge base (Agboola, et al., 2013). The fuzzy model was used to perform the fuzzification and defuzzification operations. The outputs the model predicted were compared with actual rainfall data. After simulation, it was shown that the results which were predicted were in a good sync with the measured data.

Hotel Booking Prediction by Means of Fuzzy Logic Prediction

In this research, Michael (2010) used price and web traffic to model the prediction system. These variables were later subdivided into other variables. Data was collected over a period of

three years on a monthly basis. The following variables were contained in the model:

- Percentage of month average price over the three years
- Percentage of month average year visits
- Percentage of month average year bounce rate (ratio between the number of times a visitor of the site immediately leaves the site after entering and the number of times the site is visited)
- Percentage of month average year new visits
- Percentage of month average year time on site
- Percentage of month average pages per visit

The above variables are the input variables which were used to predict the output variable "percentage of month average reservations over the three years".

3.0 METHODOLOGY

3.1 Proposed Fuzzy Logic Model

The proposed fuzzy logic model uses the Mamdani fuzzy logic approach. The Mamdani approach was chosen because it is widely accepted for capturing expert knowledge, flexible, uses defuzzification technique, and has output membership functions.

Fuzzification is the technique of transforming crisp values into grades of membership for linguistic terms of fuzzy sets. The membership function is used to connect a grade to each linguistic term (Straszecka, 2003; Nagarajan and Thyagharajan, 2004; Ibrahim and Elnady, 2013). Fuzzification of the data is done on the transformed data by entering the selected input parameter into the horizontal axis and projecting vertically to the upper boundary of the membership function to determine the degree of membership.

In this paper, *triangular fuzzy number* (TFN) is employed. The membership function μ_x(TFN) is defined below in equation (1) by (Foley and McGrory, 2011).

$$\mu_{x(TFN)} = \begin{cases} 0 \ if \ x \leq l \ or \ x > k \\ \frac{x-l}{n-l} \ if \ l < x \leq n \\ \frac{k-x}{k-n} \ if \ n < x \leq k \end{cases} \qquad (1)$$

where:*l* is taken to be lower boundary; n is median point and *k* is upper boundary, n≠l, k≠n;

Defuzzification is quite optional and suitable only when there is a need to convert the derived fuzzy number to the numeric data in the real world domains. That is the process by which a non-fuzzy (crisp) output is obtained from the fuzzy set.

The Mamdani model in this paper was designed based on various potential breast cancer risk factors. Some potential breast cancer risk factors include:

3.1.1 Age

The risk of developing breast cancer increases with age. It has been found that the risk rates are generally low in women younger than 40. This risk_rate increases after age 40 and are highest in age 70 and above.

3.1.2 Sex

Majority of breast cancer cases are normally associated with females. Breast cancer also occurs in males but it is very rare.

3.1.3 Family History

Women who have relations like a mother or sister with breast cancer have a greater risk of getting breast cancer. The risk is multiplied if the woman has multiple relatives who have the disease.

3.1.4 Alcohol

Women who take less than one drink a day have a lower risk of being diagnosed with breast cancer (one drink can be a glass of wine, can of beer, or a shot of hard liquor). It is possible for alcohol to raise the level of hormones in the body. Increased level of hormones after menopause in a woman can cause cells in the breast to become cancerous.

3.1.5 Menarche Age

Early menstruation in a woman can increase the risk of breast cancer. This is possible because an early period exposes a woman's body to greater amounts of estrogen (reproductive hormone in a woman) over her lifetime. Increased levels of estrogen in a woman over a long period of time can increase the risk that cells in the breast will become cancerous.

3.1.6 Age at Menopause

Women who have menopause at a later age have higher risk of being diagnosed with breast cancer.

This is possible because an early period exposes a woman's body to greater amounts of estrogen (reproductive hormone in a woman) over her lifetime. Increased levels of estrogen in a woman over a long period of time can increase the risk that cells in the breast will become cancerous.

3.1.7 Number of Births

Women who have less than 2 children have a higher risk of breast cancer. The reason is that pregnancy changes hormone levels and breast tissue in a way that helps protect breast cells from becoming cancerous. The more pregnancy a woman has, the more her breast tissue changes and the more protection she gets.

3.1.8 Birth Weight

A woman who weighs more when she gives birth has a higher risk of breast cancer before menopause.

Researchers are not sure yet why this occurs but studies are being conducted on different influences like pregnancy hormones and other prenatal factors to learn more about breast cancer development.

3.1.9 Tumor Size

Tumors occur as a result of formation of mass tissue which is caused by the continuous splitting and dividing of cells without control thus creating excess cells which are not needed. Tumor size measured in centimeters (cm) ranges from 0.1to 5.0cm. Women with tumor size < 1.0cm have lower risk than women with tumor size >= 5.0cm.

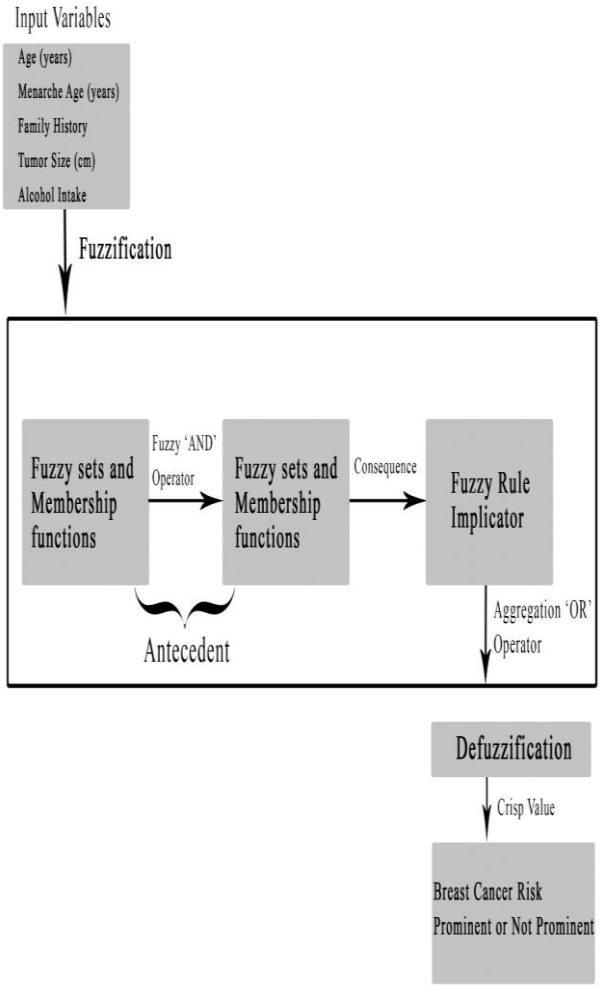


Figure 2.0: Architecture of the Breast Cancer Fuzzy Inference System

After reviewing the above potential risk factors, a few factors were chosen to be incorporated in the model. These factors are: age, menarche age, family history, alcohol intake, and tumor size. In accordance to these chosen factors, membership functions were created using the Triangular Membership Function. Figure 2. shows the architecture of the breast cancer system which involves the input variables, membership functions, rules and the defuzzification process which produces the crisp value.

3.2 Linguistic Variables

The Linguistic labels used to represent each parameter are represented in Table 3.0.

Table 3.0: Linguistic Labels for Fuzzy Variables

No	Parameters	Linguistic labels
1	Age	Range one(R1), range two(R2), range three(R3), range four(R4)

2	Menarche Age	Early(EL), normal(NM), late(LT)
3	Family History	YES, NO
4	Tumor Size	Small(SM), medium(MD), large(LG)
5	Alcohol Intake	Less(LS), more(MR)
6	Breast Cancer Risk	Low risk(LR), risky(RS), high risk(HR)

Table 3.1 shows the numerical variation interval for input variables.

Table 3.1: Numerical Variation Interval for Input Variables

Input Variables	Min Value	Max Value
Age	40	100
Menarche Age	10	15
Family History	0.1	1.0
Tumor size	0.1	5.0
Alcohol intake	0.1	1.0

Table 3.2 shows the numerical variation interval for output variables.

Table 3.2: Numerical Variation Interval for Output Variables

Output Variables	Min Value	Max Value
Breast cancer risk	0%	100%

Table 3.3 shows the Linguistic labels and the range used for each of the parameter.

Table 3.3: Linguistic Labels and Their Ranges

No	Parameters	Range
1	Age	Range one: 40 – 49 Range two: 50 - 59 Range three: 60 - 100
2	Menarche age	Early: 10 – 11 Normal: 12 - 13 Late: 14 - 15
3	Family History	Yes: 0 - 0.5
		No: 0.6 - 1
4	Tumor size	Small: 0.1 - 1.0 Medium: 1.1 - 2.0 Large: 2.1 - 5.0
5	Alcohol intake	Less: 0.0 - 0.5 More: 0.6 - 1
6	Breast cancer risk	Low risk: 1 – 33.3% Risky: 33.4 – 66.6% High risk: 66.7 – 100%

An example is using the ranges of fuzzy values in table 3.3. The linguistic values in table 3.3 are mapped to their respective membership functions using the triangular membership formula which is defined in equation (1).

Low risk will output a value x, where 1≤x<33.3

$$low\ risk(\mu_x) = \begin{cases} 0\ if\ x \leq 1\ or\ x > 33.3 \\ \frac{x-1}{16.15}\ if\ 1 < x \leq 17.15 \\ \frac{33.3-x}{16.15}\ if\ 17.15 < x \leq 33.3 \end{cases} \quad (2)$$

The median point (n) = 17.15, the lower boundary (l) =1 and the upper boundary (k) = 33.3.

Risky will output a value x, where 33.4≤x<66.6

$$risky(\mu_x) = \begin{cases} 0\ if\ x \leq 33.4\ or\ x > 66.6 \\ \frac{x-33.4}{49}\ if\ 33.4 < x \leq 50 \\ \frac{66.6-x}{16.6}\ if\ 50 < x \leq 66.6 \end{cases} \quad (3)$$

The median point (n) = 50, the lower boundary (l) =33.4 and the upper boundary (k) = 66.6.

High risk will output a value x, where 66.7≤x<100

$$High\ risk(\mu_x) = \begin{cases} 0\ if\ x \leq 66.7\ or\ x > 100 \\ \frac{x-66.7}{16.65}\ if\ 66.7 < x \leq 83.35 \\ \frac{33.3-x}{16.65}\ if\ 83.35 < x \leq 100 \end{cases} \quad (4)$$

The median point (n) = 83.35, the lower boundary (l) = 66.7 and the upper boundary (k) = 100.

3.3 Rule Base

The rule-based used for the proposed system is depicted in Table 3.4.

Table 3.4: Defined Rules for the Fuzzy Model

Rule number	Age	Tumor size	Menarche age	Alcohol intake	Family History	Output: breast cancer risk
1	R1	SM	EL	LS	YES	LR
2	R1	SM	EL	LS	NO	LR
3	R1	SM	EL	MR	YES	RS
4	R1	SM	EL	MR	NO	LR
5	R1	SM	NM	LS	YES	LR
6	R1	SM	NM	LS	NO	LR
7	R1	SM	NM	MR	YES	RS
8	R1	SM	NM	MR	NO	LR
9	R1	SM	LT	LS	YES	LR
10	R1	SM	LT	LS	NO	LR
11	R1	SM	LT	MR	YES	RS
12	R1	SM	LT	MR	NO	LR
13	R1	MD	EL	LS	YES	HR
14	R1	MD	EL	LS	NO	RS
15	R1	MD	EL	MR	YES	HR
16	R1	MD	EL	MR	NO	HR
17	R1	MD	NM	LS	YES	RS
18	R1	MD	NM	LS	NO	RS
19	R1	MD	NM	MR	YES	HR
20	R1	MD	NM	MR	NO	RS
…	…	…	…	…	…	…
60	R2	MD	LT	MR	NO	RS
61	R2	LG	EL	LS	YES	HR
62	R2	LG	EL	LS	NO	HR
63	R2	LG	EL	MR	YES	HR
64	R2	LG	EL	MR	NO	HR
65	R2	LG	NM	LS	YES	HR
66	R2	LG	NM	LS	NO	RS
67	R2	LG	NM	MR	YES	HR
68	R2	LG	NM	MR	NO	HR
69	R2	LG	LT	LS	YES	HR
70	R2	LG	LT	LS	NO	RS
71	R2	LG	LT	MR	YES	HR
72	R2	LG	LT	MR	NO	HR
73	R3	SM	EL	LS	YES	RS
74	R3	SM	EL	LS	NO	RS
75	R3	SM	EL	MR	YES	HR

76	R3	SM	EL	MR	NO	RS
77	R3	SM	NM	LS	YES	RS
78	R3	SM	NM	LS	NO	RS
79	R3	SM	NM	MR	YES	HR
80	R3	SM	NM	MR	NO	RS

4.0 RESULTS AND DISCUSSIONS

The proposed Breast Cancer Prediction System was developed using the Visual C# programming language and the chosen software development kit is Microsoft Visual Studio 2012. C# is a programming language that is designed for building a variety of applications that run on the .NET framework.

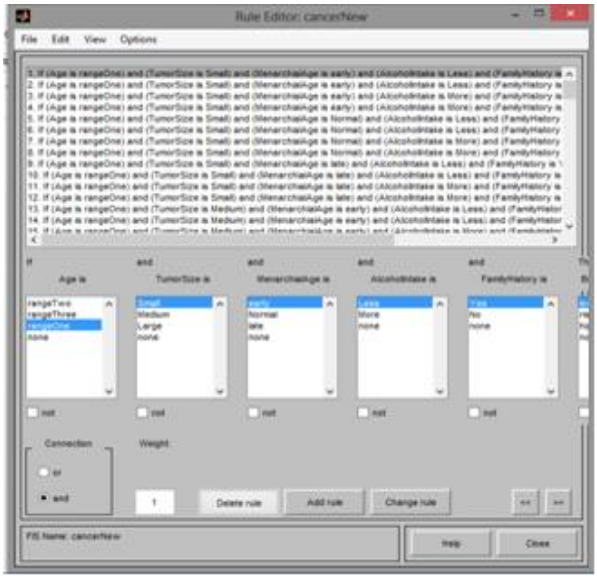


Figure 3.0: FUZZY MODEL RULE EDITOR

From figure 3.0, the fuzzy rule model interpretes such that if (Age is rangeOne) and (TumorSize is Small) and (MenarchialAge is early) and (AlcoholIntake is Less) and (FamilyHistory is No) then (Breast_Cancer_risk is lowRisk)"); and also if (Age is rangeOne) and (TumorSize is Small) and (MenarchialAge is Normal) and (AlcoholIntake is More) and (FamilyHistory is Yes) then (Breast_Cancer_risk is risky)").

This section presents the interfaces of the BCRES (Breast Cancer Risk Expert System) application.

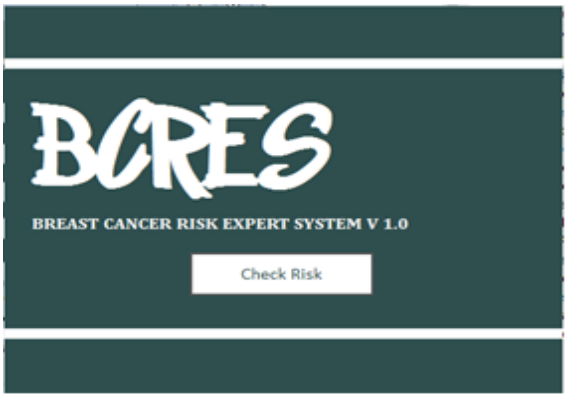


Figure 4.0: START UP FORM

Figure 4.0 describes the Start Form that is first shown to the user when he/she opens the application. The form acts like a splash screen. When the user clicks the 'check risk' button, the user is then redirected to the main form which is the risk analysis form.

The Risk Analysis Form is the form that is used to determine the breast cancer risk of a patient. The form is powered using a fuzzy logic library. Users are required to fill in all fields appropriately.

After all fields have been filled appropriately, the user is then required to click the 'Get Risk' button which shows the user the risk value of the patient being analyzed as shown in Figure 5.0.

Figure 5.0: RISK ANALYSIS FORM

The system uses the inputs such as the age of the individual which starts from 40 years and above, tumor size which ranges from 0.0 to 5.0, menarche which is a sign of a menstrual period of a woman which starts from 10 to 15. Furthermore, alcohol intake of the person and lastly, the presence or absence of any sign of breast cancer in the family background. All these inputs are matched and the risks are displayed as shown in figure 6.0. Figure 6.0 shows the result of a person who has a risk of 57.9% after all the inputs are matched and the result shows that the person has a higher risk of breast cancer. This is validated from the results collected from an expert which states that low risk starts from 1 to 33%, risky ranges from 33.4 to 66.6% and high risk ranges from 66.7 to 100%.The system was tested and validated with few individuals while some showed traces of breast cancer, the other ones showed lower risk of the ailment. One of the individuals is shown in figure 6.0

Figure 6.0: RISK ANALYSIS FORM TEST

5.0 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion

Breast cancer has been discovered as one of the most popular cause of cancer death in the whole world today. Hence, early detection of breast cancer is one of the most important means or tools for successful treatments. In this research work, a desktop based medical expert system using the Mamdani fuzzy logic approach for standardized prediction has been developed and implemented to predict the presence of breast cancer in a woman. The system alongside other medical examinations and tests would help medical practitioners predict the presence of breast cancer.

5.2 Recommendation

This research work has been designed to help its users to predict the future occurrence of breast cancer in women but it is still dependent on some other medical examinations. The fuzzy logic model used in this research work utilizes only a few risk factors for its prediction. The system can be improved further by considering other relevant risk factors which can be beneficial to the improvement of the system. Other models for prediction such as Artificial Neuro-Fuzzy Inference System (ANFIS) can also be used to enhance this research work.

REFERENCES

Abdusselam, A., Mehmet, O., & Mehmet, C. (2005). Water Consumption Prediction of Istanbul City by Using Fuzzy logic Approach. *Water Resources Management*, 19, 641-645. doi:10.1007/s11269-005-7371-1

Agboola, A., Gabriel, A., Aliyu, E., & Alese, B. (2013). Development of a fuzzy logic based rainfall prediction model. *International Journal of Engineering and Technology*, 427-435.

Artem, S. (2006). Bening vs. Malignant: Definition, Characteristics and Differences, Ontario, School of public Health.

Crimmins, E., Hayward, M., & Saito, Y. (1994). Changing Mortality and Morbidity Rates and Health Status and Life Expectancy of the Older Population. *Demography*, 31(1), 159-175.

Fatimah, A. (2009). Epidemology and Incidence of Common Cancers in Nigeria, Cancer Reg & Epid wkshop, College of Medicine, University of Lagos, Nigeria.

Foley, M. & McGrory, J. (2011). The Application of Fuzzy Logic in Determining Linguistic Rules and Associative Membership Functions for the Control of a Manufacturing Process, Masters Dissertation, Dublin Institute of Technology.

Gail M.H, Costanino JP, Bryant J, Croyle R, Freedman L, Helzlsouer K (n.d), Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer, *Journal of Natl Cancer Institute*, 91:1371-88.

Globocan (2012), Estimated Cancer Incidence and Prevalence Worldwide in 2012, Retrived from www.globocan.iarc.fr, Dec 4, 2015.

Hellman, M. (2001). Fuzzy Logic Introduction. Retrieved from www.kau.edu.sa/Files/0052079/Subjects/fuzzy.pdf

Ibrahim, H.E.A., & Elnady, M.A. (2013). A Comparative Study of PID, Fuzzy, Fuzzy-PID, PSO-PID, PSO-Fuzzy, and PSO-Fuzzy-PID Controllers for Speed Control of DC Motor Drive. *International Review of Automatic Control (IREACO), 6*(4), 393-403.

Justin, W., & Eric, Z. (2006). Prediction Markets in Theory and Practice. *National Bureau of Economic Reseacrh*, 11-12.

Justin, W., & Eric, Z. (2006). Prediction Markets in Theory and Practice. National Bureau of Economic Reseacrh, 11-12.

Malcom, R. (2001). Cancer. *Encyclopedia of Life Sciences*. Retrieved from www.els.net, March 4, 2016.

Marco Alfonse, Mostafa M. Aref, Abdel-Badeeh M. Salem (2014), An Ontology-Based System for Cancer Diseases Knowledge Management, *I.J. Information Engineering and Electronic Business*, 2014, 6, 55-63.

Michael, H. (2010). Hotel Booking Prediction by Means of Fuzzy Logic Prediction. BMI Paper.

Nagarajan, G., & Thyagharajan, K.K. (2014). Rule-Based Semantic Content Extraction in Image using Fuzzy Ontology. *International Review on Computers and Software (IRECOS), 9*(2), 266-277

Navjot, K., Navleen, S., & Anand, N. (2013). Review of Expert Systems Based on Fuzzy Logic. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(3), 1334-1339.

Ramani R. Vanitha S. S. Valarmathy S. (2013), The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images, *I.J. Image, Graphics and Signal Processing*, 2013, 5, 47-54.

Straszecka, E. (2003). Building membership functions for medical knowledge representation. *Journal of Applied Computer Science, 11*(2), 55-66.

Vaidehi, V., Monica, S., Mohamed, S., Deepika, M., & Sangeetha, S. (2008). A Prediction System Based on Fuzzy Logic, Proceedings of the World Congress on Engineering and Computer Science, October 22 - 24, 2008, San Francisco, USA

Yilmaz, A., & Ayan, K. (2011). Cancer Risk Analysis by Fuzzy Logic Approach and Performance Status of the Model, Turk J Elec Eng & Comp Sci, 21: 897 – 912.

Zadeh, L. (1965). Fuzzy Sets, Information and Control, Information and Control, Elsevier, Volume 8, Issue 3, Page 338-353.

Zadeh, L. (1994). Preface in Fuzzy Logic Technology and Applications. IEEE Software, November issue, 48-56, 1994.

Zadeh, L. (2004). Fuzzy Logic Systems: Origin, Concepts and Trends, Information Sciences, 172, 1-40.

A Review on Smart Cities: Impact of Technology and Social Factors

Mpantor Aribilosho[i] and Abel Usoro

School of Engineering and Science
University of the West of Scotland (UWS)
United Kingdom
E-mail: mpantor.aribilosho@uws.ac.uk and
abel.usoro@uws.ac.uk

Abstract: With the application of various smart technologies, industrial changes have taken place. This paper presents how improvements in technology widely change cities with special analysis of its effect on communication, infrastructure, networking, transportation, education and populace, It is proposed that research work in the future should investigate new strategies and policies that can bring about successful implementation of smart city technologies, putting into consideration the human factor, governance and technology required to create a positive change in making a city smart compliant.

Keywords*: Smart, Cities, Technology*

1.0. INTRODUCTION

Technology is and has become a global trend. The terms "smart" and "intelligent" are used as tools to market new products, nonetheless much thought is not always given as to what these really mean. A city where available public functions are networked by means of cutting-edge information and communication technologies is termed a smart city. Functions necessary for the realization of ICT-based smart city include concepts of sensor-based global environment and intelligence (Seong-Hoon and Dong-Woo, 2015). Smart city started as Internet-based virtual city in computers and has progressed to the present status. The framework of smart city consists of the six elements: smart government, smart building, smart mobility, smart energy, environment and smart service. Various products might include the topmost of great technology as available today but the defining question is, are they really capable of bringing to the fore a consciousness of their position and do they possess the required ability of responding to this awareness? Attributes like these are significant and required for an existing technology to be categorized as truly smart.

Smart people, smart governance, smart economies, smart homes, smart environments and smart devices are a few of the numerous component dimensions that are often associated with the concept of a smart city as outlined by Madakam et.al (2014). These know-hows are orchestrated by intelligent inventions which possess capabilities that allow for the sensing of variations in their environments with abilities to implement measures that improve their functionalities with respect to a change in these circumstances and according to Worden et.al (2003) these inventions massively aid proficiency, performance, productivity and the economics of operations.

Arising from the need to understand the depth and breadth of the concept of smart cities development and acceptance, it has become necessary to gain a significant appreciation of the elements that combine to make such urban developments. It is also equally important to understand the roles played by these individual elements with the aim of extracting appreciation on the subject matter.

Smart technologies are intended to be engrafted into every aspect of our lives as human beings since these technologies are aimed to make structures like sanitation, social amenities, transportation, housing, land use and land reclamation better monitored, more maintained and more economical. The Internet of Things (IoT) was concerted to enable structural entities with individual responsibilities out of the most mundane things and enable humans to access these data and take decisions from the comfort of their smart gadgets without necessarily visiting the

control centre as the case was in the recent past. Amazingly, SMART is an acronym which means Self-Monitoring Analysis and Reporting Technology. What this really means is object embedded with the ability to record data, analyze them, and make a decision based on either a programmed response, or the most preferably used options from its database of actions. The smartness of these devices could also be applied to the ability to make intelligent decisions from the information that has been imputed into them by the user or from the information gotten from the internet. These devices are capable of anticipating occurrences before they happen and take the predetermined precautions. The history of smart technology could be traced back to IBM, when they created software that helped predict when their hard drives were going to fail. The capability of smart devices has since transcended to possess high level capacities ranging from simple tasks such as lighting bulbs at specified times, to keeping space craft in orbit of intended travel.

Smart City: There are various ways a smart city can be defined. It could mean a city that has smart and user-friendly technology entrenched through all city functions; a city where all sectors are being driven by technology.

Smart city could also suggest sustainability of a population in the areas of economic activities and employment, irrespective of their education, abilities or revenue levels.

Figure 1: An illustration of a typical smart city

Source: http://blog.ideas4all.com/es/files/2013/01/Smart-city.jpg

The concept of a smart city is active and on-going, not static. Smart cities can be described as a process by which cities become better places to live in with favourable and user-friendly conditions. It can also be described as a city that identifies the physical restrictions to its advancement without affecting the value of life of the present–day and upcoming generations.

Technology: Technology includes products or methods that make life more enjoyable and stress free for the people. A typical smart city has products of technology like: Wi-Fi, video surveillance, traffic control lights, electric street lights, and lots more which distinguish the smart cities from the local countryside. The Smart City Technology impacts on the following aspects of the environment:

- Communication and Networking
- Infrastructural Design
- Economy and businesses
- Transportation and Landscape
- Education

2.0 AIM: The purpose of this paper is to investigate how Smart Technologies create a reliable change in the society and how it transforms the health, socio-economic life, business, education, transport, energy, landscape sectors, and other diverse aspects of modern society. It also aims to propose areas for further research.

3.0. THEORY REVIEW

As identified by Frost & Sullivan (2014**)** there are eight important features of a smart city: smart energy, smart building, smart governance, smart infrastructure, smart healthcare, smart mobility, smart citizen and smart technology.

New technology has impacted businesses in three major ways which are defense, efficiency and winning. In defense, when the advancement in technology is ignored, the impact is that one loses. It was opined that there are no choices but to carry on. Another factor is that of efficiency. Functionally, technology saves costs, it produces efficiencies, and it delivers better communication.

Technology gives you a head start in dealing with your clients and customers as well as your products and services. The third aspect is winning. You can use technology to distinguish yourself, and make your products better than that of competitors.

The five cities on the verge of achieving the dream of becoming smart cities include: London, Bristol, Manchester, Glasgow and Birmingham. These cities are achieving this by relative interconnectedness of various parts of the cities ranging from door locks to power grids (Doorn, 2014**)**. These cities have become actively engaged in intelligent maneuvers such as collating data that are inputted by the diverse types of sensors that have been embedded inside them. Of the many possible applications that the interaction enabled platforms that smart technology could provide, the most evolved and closest to reality are itemized below:

a). Intelligent Transportation Systems (ITS) includes a wide variety of digital technologies which help to develop better efficiency in the transport sector. The use of ITS on vehicles, public transportation network and roads can surely aid reduction in traffic congestions, improve mobility, save lives, and optimize other existing infrastructures. This innovation is gradually changing our outlook and adding a new dimension to travelling which includes but not limited to:

- Providing end users with parking space information, real-time traffic locations, transit and navigation info;
- Improving highway and vehicle safety;
- Minimizing traffic congestion and incident response times;
- Reducing energy emissions and use;
- Improving supply chain efficiency and shipment;
- Enabling advanced financing substitutes such as dynamic pricing and mileage-based user fees. (smart Transportation Technology Showcased on Capitol Hill held in June 26, 2013)

b) With the outcry by the UN on global warming, many cities are striving hard to transfer their insatiable energy demands from the exhaustible fossil fuels to more environmental and greener sources of energy. Smart technology designed with the goal to study and proffer the most efficient way to achieve this based on the weather data could be engaged to tackle this disturbing challenge as seen in the San José Green vision being implemented in California, United States. Intel and San Jose are creating a "sustainability lens" by installing a network of air quality, sound and microclimate sensors. The "sustainability lens" uses IoT (Internet of Things) technology to measure characteristics such as particulates in the air, noise pollution, and traffic flow. With this information, the city management can work on air quality improvements, noise, efficiency in transportation, health, energy and also environmental sustainability (see www.smartamerica.org).

The goals to be achieved via this program are surmised below:

- Charging stations for electric vehicles conveniently located just off freeways.
- Smart Meters in homes and businesses that regulate energy consumption throughout the day
- Photovoltaic solar panels on every roof allowing owners to independently produce energy.
- Recycled rain water containers on every roof that provide water for domestic use, ensuring water conversion
- Personalized digital signs that provide relevant and personalized information for people passing by
- M2M (motorist-to-motorist) technology that provides drivers with up to the moment information about traffic delays and the most effective routes.
- Biometric technology that confirms identity and allows users to open buildings, vehicles, computers and other secure areas with a simple touch of their fingertip
- Super speed train systems offering reliable wireless Long Term Evolution (LTE) 4G

access providing users with constant access to networks anywhere
- Efficient bicycle lanes to assist the growing number of zero-emission vehicle users

c) The ever growing abuse of weapons and gun-related violence in major cities has risen to catastrophic heights in recent times. A way to curb this resulted in an innovation in Europe by Armatix, a German company which created a line of smart guns which can only be fired by authorized personnel whose biometric data are accessed before the security protocols of the gun could be unlocked. Another innovation is the use of unmanned drones by the US government in military expenditures as a tool for reconnaissance which performed well without the unnecessary loss of human life. This ever evolving and highly secretive technology are programmed to make cities safer. The civilian end of this smart security technology includes:

- Burglary protection, which uses smart alarms and auto-call features.
- Home automation technology, which uses individual living patterns to detect suspicious developments.
- Fire and carbon monoxide monitors which look out for rise in the level of smoke and CO_2 gas before they rise to toxic proportions.
- In case of accidents, at home or the road, the cars or the house would make emergency calls to the necessary authorities (Smart home services, 2014).

d) The use of smart technology in health care cannot be over emphasized. The devices such as pacemakers and dialysis machine have been very efficient in their jobs. The use of technology has helped to reduce the number of human-related dangers posed in many of these disciplines. There are no limitations to possibilities which eventually encompass food production in agriculture and production lines.

4.0. THEORETICAL FRAMEWORK

4.1 INFRASTRUCTURE DESIGN

The way cities are designed and operated have always been bordered on the transformative effect of technology. Spreading suburban areas were stimulated by the invention of automobiles; the invention of elevator resulted in multistory and skyscraper developments; a global economy nurtured in municipal clusters is possible due to shipping and affordable air travel; also infrastructural developments including networks of public transport and systems managing waste collection and furnished mega cities of huge populations with the motivation to continue growing.

In times past, local constructions have been driven by issues such as sanitation, fire or flooding. The reaction has been to engineer systems (under the powers of centralized, state-led planning and public funding) that solve a single problem at a particular time. Not much consideration has been given to future conditions. Nowadays, and especially with the advent of Information and Communication Technology (ICT), a major wave of development with respect to technology is sweeping through cities resulting in developmental changes, inspiring new city experiences and stimulating their functionality.

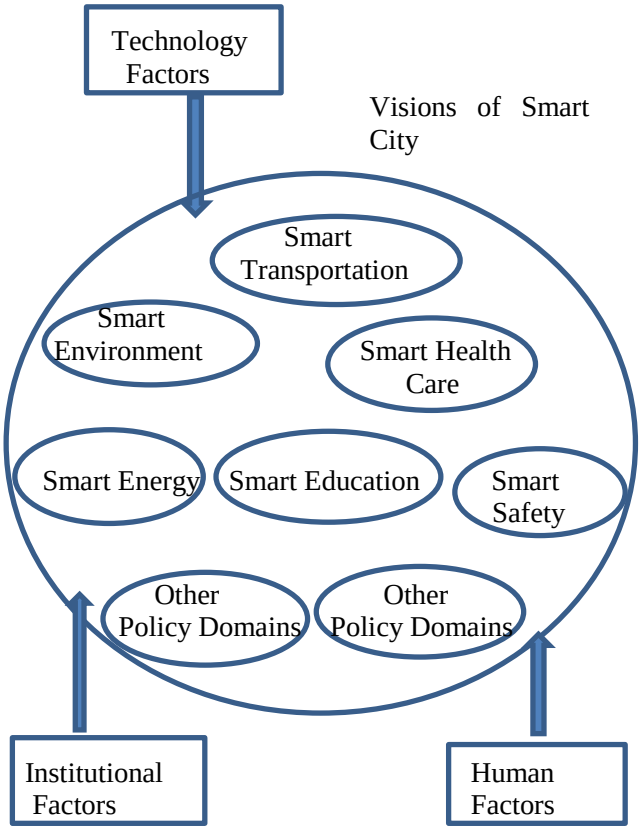


Figure 2: Directions of Smart City, figure adapted from Nam and Pardo (2011, p. 288)

4.2 TRANSPORTATION AND LANDSCAPE

A city cannot be called a smart technology city if it does not have digitalized road networks with street lights, Wi-Fi cameras, surveillance video and the like.

Movement supports all we do individually; as groups and global economies.

Individuals want to move about to achieve simple social necessities, but movement is an amenity, adding to value of life by allowing study, relaxation and reformation. High value movement in the city is essential for the realization of other municipal sectors and the establishment of jobs, and acts a key part in refining a good-looking environ for inhabitants and businesses.

Hitherto mobility is extensively referred to as one of the greatest inflexible and global challenge tackled by cities worldwide. With an increase in urban populations, present and developing metropolises face the task of meeting growing needs for effective mobility in unfinished physical groundwork capacity. Concurrently, inhabitants' hopes are fluctuating recurrently, due to continuing modernisms around low carbon and proficient means of transportation technologies and developments in infrastructure organization. (Arup 2010).Smart cities now employ more sophisticated technologies in the planning and development of settlements with such considerations as traffic flow, pedestrian access, and commercial heavy duty trucks and machineries movement in consideration from the basic design stage. This allows for the incorporation of required sensors and aiding devices to be installed as at when due with minimal cost implications.

4.3 EDUCATION

The advent of the Internet brought with it the possibility of expansion on learning from a distance which has been in practice for decades now. Expansion in terms of coverage and better delivery of instruction as a result of Information and Communication Technology (ICT) as compared to the days of old. Texts are now easily combined with audio and video pieces of information for the purpose of maximizing time and getting the best out of education. Teachers and students in smart cities now have the ease of interacting via social media, e-mail and discussion groups. Teachers and students alike readily have access to education audiovisuals online anytime and irrespective of location. A teacher in a smart city can teach a student in another smart city over the internet; this would be impossible with traditional teaching methods.

The essence of distance learning via the Internet is to complete and complement teaching methods that are already existent instead of competing with these traditional methods of learning as existed before ICT.

A smart city is one that would capitalize on the availability of technology and direct it for the betterment of education processes which is the foundation for technology and innovation. Resulting from the advantages of access to internet and the suitability for its use in learning as well as the convenience of learning from a remote location, there is expected to be a rise in the demand for distance learning programs which will subsequently metamorphose to development and advance in the commerce of education.

4.4 POPULACE

Towards advanced smart cities, the starting point should be from human assets, instead of just trusting that IT itself can spontaneously change and develop cities (Hollands, (2008). The serious factor in any efficacious city is its people and how they interrelate. Sturdier approaches to education, responsiveness, and control offer services that are reachable to all, barriers associated with language, culture, education, disabilities, skills and development should be gotten rid of (Coe et al, (2001). For those who might be lagging behind with respect to the use of latest technologies, social learning addresses this concern. Teaching and training activities should improve IT skills, cultivate information workers, enable the environs of social learning, and develop IT training in institutes, societies and businesses (Cairney & Speak, (2000).

4.5 COMMUNICATION AND NETWORKING

Owing to continuous advancements in the smart phone technology which focuses on its unceasing and more reliable connection to the internet, processing and storage functionality improvements, better computing abilities, advancements in sensing and modeling capabilities and the Internet of Things (IoT), the economics of cities and the dynamic nature of urban settlements are being reshaped in a collective fashion. These machineries symbolize an important prospect for cities and their administrations to craft more effective, operational and promising urban settlements.

The ability to converse in ways similar to traditional face to face conversations is made available by computers and associated communication technologies. Smart technologies allow for collaboration between people who do not have to meet in person. This suggests that there are plenty of activities including, but not limited to, conferencing, teaching, learning, attending a function, judging a case, and many more that can be done by smart proxy, instead of in person saving time and resources that might otherwise have been spent.

4.5.1. COMMUNICATION: SMART PHONES IN SMART CITIES

The concept of computing, amalgamated with or infused in communication for consumers who use mobile smart technologies like tablets and mobile phones, is on a progression path, directed towards achieving interoperability with its corresponding influence on functional facilities from each and every industry.

Taking the lead in this union are smart phone devices which play an ever important role as the universal mobile terminal. This process of convergence is made possible by the incorporation of smart phones which can be defined in real sense as a mobile phone that performs many of the functions of a computer due to its advanced operating system and features that enables possibilities beyond traditional functionalities like calling or texting. These devices are a common sight in every smart city and serve as a handy substratum for the interphase with applications, devices and systems within a smart city.

There is also the concept of smart phones functioning with existing platforms to enable users to communicate with systems in such a manner as to derive convenience and save costs. Hence the communication between these smart mobile devices and sensors or applications could be helpful for example in avoiding real-time traffic or an overloaded bus. Also smart devices could be used as sensors to share information gathered by the device for the purpose of improving a system. This act of sharing termed as "participatory sensing" ensures communication between the developers of a system and the end users with the information gathered being used for improving city services.

On the flip side, as a marketing strategy, smartphones of various specifications and functions are becoming increasingly affordable. This comes with its corresponding nuisance effect and could be seen by many city dwellers as a distraction. The habitual tendencies associated with the use of smart phones could result in lack of awareness and be detrimental to user's safety. In these circumstances, it is arguable that a smart city can be a dangerous city. Our smart phones, when combined with the infrastructures and technology available within the cities, could have the ability to unquestionably change our world. The big and debatable question is: what is the direction of change?

4.6 BIOROBOTICS

Many claim that the challenge that is faced in urban cities is inadequate investment in robotics. Robotics generally involves the use of a computer aided robot to perform specific tasks as designed. Industrial robots are reputable for repetitive and routine tasks mostly in mass production facilities. Service robot development is an area of robotics yet to be maximized.

With some connectivity and innovation robots are capable of making our cities smarter than they currently are. Unlike robots in the industrial world, service robots are designed to carry out specific tasks that would normally be taking place in an unstructured setting with the likelihood of direct human-robot interface. There is the concept of co-worker robots, designed to assist human beings in carrying out a number of different tasks but this has not found full use in numerous countries

owing to the expenses associated with owning and operating them and the fear of resulting unemployment due to robots taking human jobs.

Whereas the prospects of having robots assist with routine tasks within the home are very promising, their use in cleaning floors has been the most successful application, especially in Korea.

The lack of a universally accepted standard as well as costs and acceptability in some opinion is responsible for the slow pace of development in these fields.

5.0. CONCLUSION

Smart technologies in cities function to make existing structures more efficient. Since the aim of smart technologies is to eliminate the gargantuan problem and the grotesque odds facing survival in the city life, a city would only be referred to as "smart" if there are technologies in place that amongst other things, help people move around freely, make services and amenities identifiable and reachable, be aware of happenings around with the aim of staying safe and most importantly be readily accessible. This progress is not without its oddities, the easier it is for the individual to survive in the city, and the quicker it is for such city to be over-populated, which in itself is a challenge not easily managed. With smart cities comes the urge to urbanize indiscriminately which is also another huge challenge for humanity. Smart cities open the possibility to commit smarter crimes by personalities who understand the sophistication required to break many of the security codes. Another pitfall is that a minute computer error could be a devastating moment for residents of such cities. An example was when an American communication satellite went out of orbit rendering pagers useless for a long while. But as with all human achievements that has come to stay, its advantages grossly outweigh the disadvantages.

It is quite easy to see the benefits of smart cities but the question to be answered is: where do we ultimately find ourselves with the adoption of smarter cities and technologies? We imagine that the expounded concept of smart city in this write-up will add to future studies. The idea brings to the fore that there is a linking among technical, social, and institutional factors. Today the use of "smart" captures advanced and transforming changes driven by innovative technologies. Nevertheless, communal issues other than smart tools are central to smart cities. In this regard, a socio-technical opinion on smart city is required. A broad understanding of the connections between social and technical aspects of the environment is important. Research should be carried out on the roles governments, private sectors and policies play in the adoption of smart technologies as it is evident that there must be a complete buy-in for substantial developments to manifest as regards building smart cities. To this close, we will carry on with the study on smart city by concentrating on smart city creativities, bearing in mind the subtleties of several key stakeholders in those initiatives, and conferring policy improvement in urban administrations. Lastly, I would also recommend that an investigation on how smart cities affect control and freedom of citizens be carried out for a richer understanding of the concept.

REFERENCES

Armatix [online]. Available: <http://www.armatix.de/?L=0>

Arup, (2010). *Smart Cities cornerstone series Urban Mobility In The Smart City Age.*

Brad Reed. (2010). A brief history of Smartphones. Available: <http://www.networkworld.com/slideshows/2010/061510-smartphone-history.html#slide1>

Business Dictionary. Available: <http://www.businessdictionary.com/definition/technology.html>

Cairney, T., & Speak, G. (2000). Developing a 'Smart City': Understanding Information Technology Capacity and Establishing an Agenda for Change. Sydney, Australia: Centre for Regional Research and Innovation, University of Western Sydney. Available at http://trevorcairney.com/file_uploads/cgilib.30886.1.IT_Audit.pdf.

Chris Ziegler. (2011). *Android: A visual history.*

City of San José Green. (*2012). Annual Report Executive Summary*

Coe, A., Paquet, G., & Roy, J. (2001). E-governance and smart communities: A social learning challenge. Social Science Computer Review, 19(1), 80-93.

ComScore, Inc.(2012) *Mobile Future in Focus.* Available: <http://www.comscore.com>
Frost & Sullivan (2014). Global Smart Cities market.
GSMA. Technology creates smart cities. Available: - <http://www.gsma.com>
Hamza Querashi. (2012). *Evolution, Features and Future Review*. Available: <http://www.thenewstribe.com/2012/07/16/apple-from-iphone-1-to-iphone-5- evolution-features-and-future-review/ (2012)>
Hollands, R. G. (2008). Will the real smart city please stand up? City, 12(3), 303-320.
Intelligent Transportation Society of America. Available: <http://www.itsa.org>
James N., & *Nancy G.*, (2010). *A Brief History of Palm*. Available: http://www.pcworld.com
Madakam, S., Ramaswamy, R., & Tripathi, S. (2014). *Internet of Things (IoT): A Literature Review. Journal of Computer and Communications* 3 (05), 164.
Menno van Doorn (2014). What smart technology actually mean. Available: <http://labs.sogeti.com>
Nurfit (2012). *Smartphone Addiction and Impact on Society*. Available: <http://nurfitriah.wordpress.com>
Pei Z., & Lionel M., (2006). *Smart Phone and Next Generation Mobile Computing*. Available: <http://www.sciencedirect.com>
Sam C., (2012). Available: <http://ipod.about.com>
Smart cities: The future of urban infrastructure. Available<www.bbc.com/future/story/20131122-smarter-cities-smarter-future>
Smart city: *How 'Smart' Tech Will Help Create A Sustainable Future*. Available: < www.wearesalt.org>
Smart Cities USA. Available: <http://smartamerica.org>
Smart home services ADT Security Services.Available: - <http://www.adt.com>
Seong-Hoon Lee & Dong-Woo Lee (2015) Review on Present Situations of Smart City. Advanced Science and Technology Letters 98, 31-33 http://dx.doi.org/10.14257/astl.2015.98.08
Taewoo, N., & Theresa, P. (2011). Conceptualizing Smart City with Dimensions of Technology, People, and Institutions. The Proceedings of the 12th Annual International Conference on Digital Government Research, 282-291.
Techterms. (2010). Available <http:/ www.techterms.com>
University College London (2014). Delivering the smart city. Available: <https://www.ucl.ac.uk>
Wikipedia. (2012). Available: <http://en.wikipedia.org>
Worden, K., Bullough, W. A., & Haywood, J. (2003). Smart Technologies.

[i] Mpantor Aribilosho is Dr Abel Usoro's first year PhD research student.