ISSN: 2572-8830 (Print)
ISSN: 2572-8849 (Online)

American Journal of Advanced Research

Volume 1, Number 1, December 2017

Published by Institution of Business Intelligence Innovation

American Journal of Advanced Research

Vol. 1, No. 1; December, 2017

DOI https:// 10.5281/zenodo.1227463

Table of Contents

Editorial Board

American Journal of Advanced Research, 2017, 1–1
Sept. 2017, pages 1-6
doi: 10.5281/zenodo.807666
http://www.ibii-us.org/Journals/AJAR/
ISBN 2572-9829 (Online), 2572-9810 (Print)

A Feature Enhancement Method for 3D Tree Synthesis

Ling Xu[1, *]

[1]Department of Computer Science and Engineering Technology, University of Houston-Downtown, USA

*Email: xul@uhd.edu

Received on May 05, 2017; revised on May 29, 2017; published on June 13, 2017

Abstract
Computer synthesized 3D tree models often appear in the virtual worlds of 3D movies and games. The creation of tree models with controllability of realistic features including branch irregularities and gaps between branch clusters is a challenging task. This paper presents a procedural method for 3D tree feature enhancement. The tree model is generated in a weighted graph and the basic geometric tree branches are least-cost paths from path planning. Through disconnecting edges in specified regions of holes in the graph, the method can control the tree branches to avoid the hole regions thus demonstrate crooked shapes. The distribution and the sizing of holes can be used to enhance the tree features of irregularity and heterogeneity.
Keywords: Procedural modeling, Tree modeling, Geometry synthesis

1 Introduction

With the fast development of computer-generated movies and games, the audiences and game players have high demands of realistic experience for the synthetic virtual scenes. Synthesizing realistic 3D worlds becomes a more and more important topic in both computer research and industrial areas. As a commonly seen component of the nature world, trees play an important role in decorating the 3D natural scenes, by attracting the viewer attention and invoking a realistic sense of beauty and complexity. An example of virtual scenes with trees is shown in Figure 1, where the high trees and small bushes decorate the environment and make it realistic. The trees have complicated and hierarchical structures, composed of a variety of branches including thick trunks, multiple primary branches, and many slim twigs. For computer modelers, the synthesis of the numerous branches is a tedious work. In addition, different species of trees have different features such as branching patterns, branch shapes, and branch distributions. Some trees, such as spruce trees (as shown in the left of Figure 2), have straight branches and side primary branches attach to a straight trunk; some trees, such as Japanese maple trees (as shown in the right of Figure 2), have multiple primary branches without a distinct central trunk and the branches are highly crooked. How to efficiently synthesize 3D trees with controllability of a variety of branch features is also a challenge for computer modelers.

This paper presents a method to synthesize 3D tree models with effective enhancement of tree features. The method is based on the idea of path planning introduced in our earlier work (Xu and Mould, 2007): a graph is built by placing graph nodes uniformly in a 2D plane or a 3D space, and edges connect nodes and are set with specified weights. A path between two nodes is defined as those successive edges that connect one node to the other, whose cost is the sum of the weights of its constituent edges,

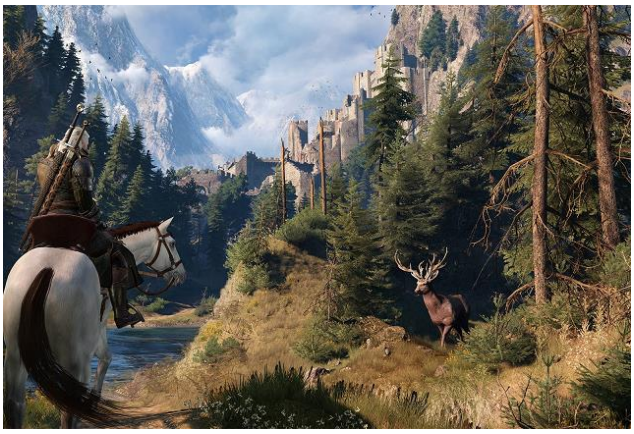


Fig. 1. A virtual world with trees in the game "*The Witcher 3*". Image from Flickr.com.

Fig. 2. Two real trees. Left: a spruce tree; right: a Japanese maple tree. Images from Flickr.com.

and the least-cost path is the one which has the minimum cost among all paths. By finding the least-cost paths connecting a node (which represents the tree root) in the lower region of the graph and multiple nodes (which represent branch ends) in the upper region of the graph, we can get a collection of least-cost paths that resemble a branching tree structure where each individual least-cost path is a tree branch. In this paper the technique is refined by involving a graph modification method, which can effectively enhance the features of resulting 3D tree models. To be specific, edges in a few regions of the weighted graph are disconnected, thus preventing the least-cost paths from crossing the regions. As a result, the synthesized tree branches will avoid the regions and demonstrate crooked shapes, making the tree appearance more irregular. In addition, the size and the distribution of the regions can contribute to create different scales of the irregularities, such as big gaps between branches and small crookedness along individual branches. These are helpful to simulate the corresponding features that can be often observed in real trees.

This paper makes the following main contributions.

- First, it proposes a method of modifying the weighted graph by disconnecting edges in specified regions. The prior methods (Xu and Mould, 2007; Xu and Mould, 2015) do not have an effective way to control the branch irregularities; the individual branches are generally smooth. The method of disconnecting edges in specified regions of a graph can effectively enhance the irregular features of branches and make the branch shape more crooked and realistic.
- Second, the exploration of varying the size and distribution of regions for disconnected edges contributes a mechanism of building different scales of branch irregularities. Large sizes of compactly placed regions contribute to create large scale of branch irregularities such as gaps between branch clusters; small sized regions contribute to build fine crooked branch shapes. This feature controlling method is novel and effective.

The paper is organized as follows. Following the introduction, the section of background will review related methods in tree modeling. The details of the feature enhancement method will be introduced in the algorithm section. Results will be given and discussed in section 4. The last section includes a conclusion and also proposes the future work.

2 Background

Due to the wide uses of tree models in computer applications such as video games and 3D movies, many methods have been devised for tree synthesis. Scanning is a common method. Based on a real tree, scanning can obtain point clouds of tree data which will then be used to reconstruct a virtual tree structure. Scanning methods (Xu et al., 2007; Li et al., 2013; Livny et al., 2010) can produce high quality tree models with precise descriptions of tree features due to the data directly from real trees. However, because of the occlusion of tree branches and leaves, scanning needs pre-processing and post-processing work for the raw data. The availability and operations of scanning devices are also not easy for novice users (Capizzi, 1982). In addition, scanning requires a real tree as an example, and it is hard to create a 3D virtual tree without an existing tree in the natural world.

Another category of widely used methods is procedural methods. Procedural methods (Ebert, 2005; Prusinkiewicz and Hanan, 1989) use algorithms with a set of parameters to depict the complex details of tree structures. By changing the rules or formulas and parameter settings, procedural methods can automatically generate a large amount of variations of results. For synthesizing complicated structures such as trees, procedural methods have their strength in the data amplification ability – "a few parameters (or a small amount of geometry) magically expand into a large, detailed model" (Ebert, 2005). A famous group of procedural methods for tree modeling is L-systems. L-systems are a grammar that generates strings that can be subsequently interpreted as branching structures such as trees. The basic idea of L-systems is based on rewriting rules. A rewriting rule is composed of strings of symbols. Each symbol represents a graphical operation, such as the following.

F Move forward a step of length d
\+ Turn left by angle δ
− Turn right by angle δ
[Push the current state onto a pushdown stack
] Pop the current state onto a pushdown stack

After applying the replacement rules for a few iterations, the initial structure can be developed into a large complex shape. An example is shown in Figure 3, where the initial vertical line (at iteration 0) can be developed into a branching structure after applying the rewriting rules for 5 iterations.

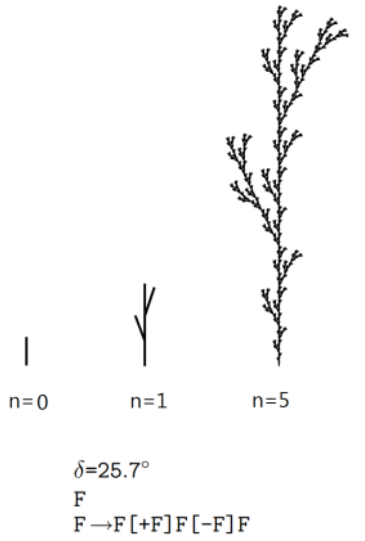


Fig. 3. Structures generated by iterations of rewriting. The example is taken from "*The Geometric Beauty of Plants*" (Prusinkiewicz and Lindenmayer, 1990)

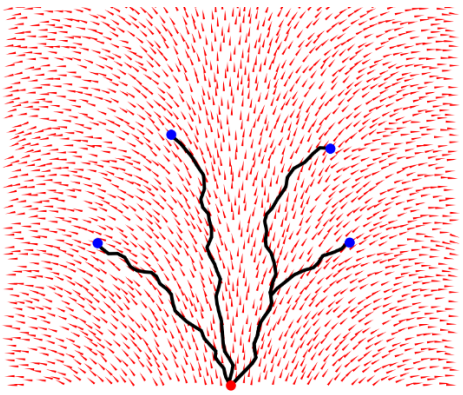


Fig. 4. Least-cost paths in a 2D graph with guiding vectors (Xu and Mould, 2015)

One issue with L-systems is the difficulty to design the rewriting rules to create a desired tree shape. It is not easy to relate the long string of symbols with the resulting tree structure.

The method introduced in this paper is built on our earlier procedural algorithms of path planning (Xu and Mould, 2007; Xu and Mould, 2015). Here we briefly review their basic ideas.

The basic tree modeling mechanism (Xu and Mould, 2007) is based on the idea of path planning. Path planning is the problem of finding the least-cost paths in a weighted graph (Millington and Funge, 2009). By selecting a node in the lower region of the graph and multiple nodes in the upper region of the graph, the least-cost (or shortest) paths connecting each upper node and the lower node form a branching structure. The purpose to select a lower node (which we call the root point) and a few upper nodes (which we call endpoints) is to simulate the tree shape – a root located at the bottom and multiple branches connecting the root and the upper ends. An example of least-cost paths is shown in Figure 4, where the red point is the root point, and the blue points are endpoints. The black paths are least-cost paths that connect the root point and the endpoints.

There are a few elements in the above path planning algorithm that can affect the resulting tree structures: the placement of graph node, the distribution of the root point and the endpoints, and the setting of edge weights. In order to make the coarse branching structure resemble real tree structures, the challenges lie in the control of branching patterns and intermediate scale architectures through the above elements, which are addressed by using guiding vectors (Xu and Mould, 2015). In the method of guiding vectors, a graph is built by connecting nodes scattered in a specified volume (e.g. a cube) with edges. Each graph node has a vector, i.e. guiding vector, used to set the weights of the edges that connect the node and its neighboring nodes. Based on a large or small deviation of the edge direction to the guiding vector direction, the weight of an edge can be expensive or cheap, thus leading to the resulting least-cost paths to follow the guiding vector directions. The least-cost paths in a graph with guiding vectors at each node are shown in Figure 4, where the red arrows are guiding vectors. The guiding vectors are set in the process of applying Dijkstra's algorithm (Dijkstra, 1959) in a brushfire way. The Dijkstra's algorithm is an algorithm to compute the least-cost paths from a source node (i.e. the root node in our case) to other nodes in a weighted graph, by proceeding from the source node to reach other graph nodes and computing the path cost from each graph node to the root node. When a graph node is visited, its guiding vector is set as a rotation from its parent node's guiding vector, leading to the smooth flow of guiding vector directions. Because the method uses the positions of endpoints to control the global tree shape and guiding vectors to vary the branching patterns, which are based on the user's geometric understanding of the target tree shape rather than abstract symbols or rules, it is straightforward to create a desired tree structure.

A complicated tree can be synthesized by adding more branches through iterations. A tree example is shown in Figure 5. The left is a tree skeleton and the right is a complete tree after adding more branches to the previous structure. The tree shows a reasonable structure and details. However, due to the setting of guiding vectors is based on an incremental rotation and the flow of guiding vectors is smooth, the branch shapes are generally smooth and lack natural irregular crookedness. Another observation of the tree model is that its branches are generally evenly distributed, however the distribution of branches in natural trees are often more heterogeneous – they have big gaps between clusters of branches. Although the guiding vectors can affect the branch shapes, making branches possess the above features through setting guiding vectors is difficult: unlike incremental rotations, making the flow of guiding vectors crooked requires a much more tedious and careful setting mechanism. In this paper, these target features are obtained by modifying the graph rather than setting guiding vectors, which can effectively address the problem. More details are introduced next.

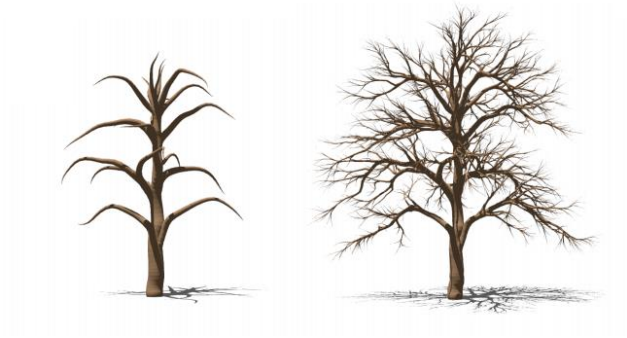


Fig. 5. A tree model created in the earlier work (Xu and Mould, 2015).

3 Algorithm

The algorithm for feature enhancement is based on the influence of edge connections to the resulting least-cost path shapes. To be specific, since a least-cost path is composed of edges in the graph, if the graph has some holes where edges are disconnected, the path between two nodes will not cross the holes and take a direct route, but tends to take available edges in the vicinity of holes. As the result, the path shape is less smooth but becomes more irregular. The idea is illustrated in Figure 6. In this figure, the left shows a graph composed of nodes (in blue) and edges (in black). A least-cost path (in red) connects two points from P_1 and P_2. On the right side, the edges that connect nodes inside the green circle and between inner nodes and outer nodes are disconnected. Because there is no edge connecting the nodes inside the hole (the green circle) and other nodes, the least-cost path that connect P_1 and P_2 will not take the original route that crosses the hole region, but takes the edges close to the hole and demonstrates a more irregular shape than the left path. Although there are other optional paths that can connect P_1 and P_2, the least-cost path tends to take the shortest route, i.e. the edges tightly follow the contour of the hole, making it feasible to control the path irregularities through the properties of holes.

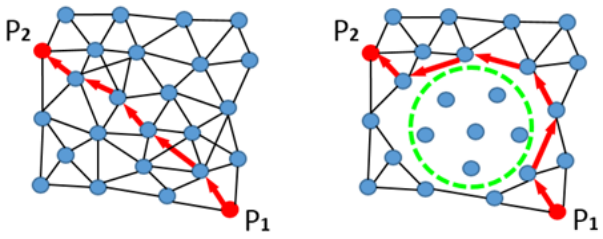


Fig. 6. Left: a generally smooth path connecting two nodes in a graph without holes; Right: a more irregular path in a graph with a hole (the green circle).

The above described algorithm involves two further questions: how to decide the placement of a hole in the graph and how to decide the size of a hole. The following introduces the explorations with respect to the above questions.

Since the least-cost path does not take in-hole edges, through controlling the placement and sizes of holes, we can change the features of resulting least-cost paths. The holes in a 2D graph (as shown in Figure 6 – 9) are defined as circles, and in a 3D graph (as those used in Figure 10 and 11) are spheres. Edges are disconnected if they locate inside a hole or cross the boundary of a hole. The purpose of doing so is to prevent the least-cost path to take the edges in the hole. The positions of holes are decided according to the Poisson Disc distribution (Fiume, 1995), where holes must keep a minimum distance to each other. This decision helps to control the distribution of holes and the shapes of the resulting least-cost paths.

The principle to decide the size and the number of holes is not that strict. Since this work does not aim at a precise simulation of a specific tree branch shape, a hole at a radius of 3 nodes will not have big difference with a radius of 4 nodes or 5 nodes. We intend to achieve a general enhancement of branch features that are absent in the previous work, so the possession of crookedness and branch gaps are the main concerns. According to our experience, depending on the target branch features, the size of holes can be set in a flexible range: in a graph of 150,000 nodes, holes at a radius of 2-6 nodes can be used for small-scale crookedness, and 10-20 nodes for large-scale crookedness. Further adjustment is possible for specific trees. Fortunately due to the benefits of procedural methods, adjusting the hole sizes does not require extra effort but changing a parameter setting. Changing the hole size and re-running the program take less than 10 seconds for most tree structures. The number of holes in the graph is decided by the minimum distance in the Poisson Disc distribution and the radii of holes. The principle to follow is to make the holes tightly and evenly filled the graph but also leave narrow non-hole regions (at a width of 2-5 nodes) for least-cost paths.

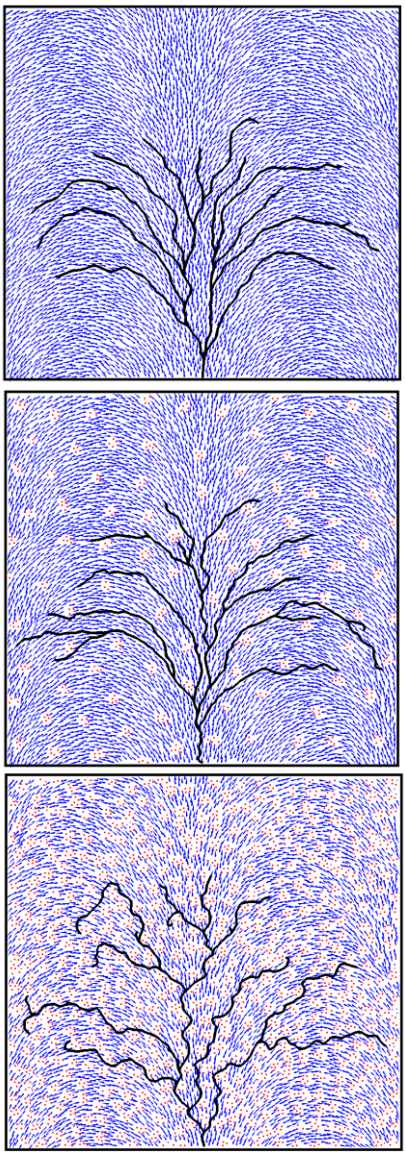


Fig. 7. Top: a tree-like structure in a graph without holes; middle: the graph has sparse holes; bottom: the graph with denser holes

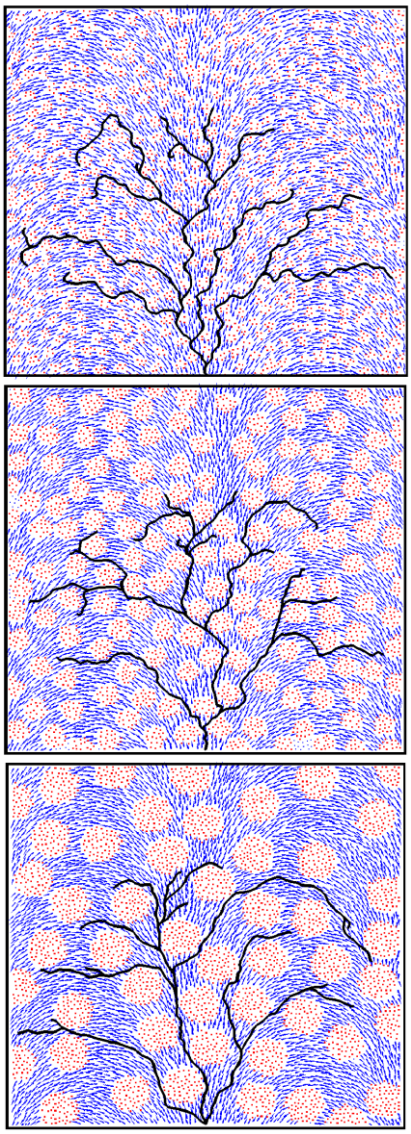


Fig. 8. From top to bottom: least-cost paths obtained in graphs with small sized holes, intermediate sized holes, and large sized holes.

In Figure 6 – 9 we use least-cost paths in a 2D graph to illustrate the algorithm for the ease of viewing the path shapes and the holes, which in 3D may be difficult because of node shading and occlusions. Figure 7 shows

how different densities of holes affect the resulting path shapes in a 2D graph. In the top, least-cost paths in a graph without holes tend to follow the flow of guiding vectors (blue arrows). In the middle, the sparsely distributed holes do not have much influence on the path shapes, because in addition to the edges close to the holes there are still plenty of other edges in the non-hole regions. It is possible that the least-cost paths do not take the edges around the holes but in a more direct route in the big non-hole regions. In the lower, the graph has many holes and the resulting paths demonstrate the crooked shapes, due to the very few available edges in the non-hole regions. A least-cost path must avoid many holes to take available edges in order to reach the endpoint, leading to the irregular path feature.

While the densely distributed holes contribute to the feature of path irregularity, the sizes of holes affect the scale of irregular features. Figure 8 shows three examples. The top graph has dense small holes in which edges are disconnected. The paths have small-scale irregularities. The holes in the middle graph are larger, leading to larger crooked path shapes. The paths in the bottom graph show large-scale features – big gaps between some paths, and large turning path shape due to the path traveling around big holes. Figure 9 shows the use of a combination of small holes and big holes. The graph in the right has 12 big holes and 240 small holes whose radii are only 1/5 of the big holes. Compared with the paths in the left graph without any holes, the right paths demonstrate both big scale irregularities such as gaps between paths and crookedness of individual paths.

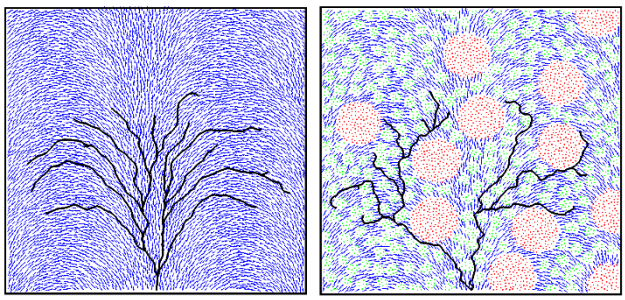


Fig. 9. A 2D tree structure with (left) and without (right) feature enhancement.

4 Results and Discussion

The feature enhancement algorithm introduced above has been applied to create 3D tree models, shown in the following figures. The evaluation of the results is based on the side by side comparison with a tree model without the feature enhancement. This comparison is built on visual inspection, which has been used for evaluating tree modeling since the earliest synthetic trees in 1980s. It is worth mentioning that in the area of tree modeling there does not exist an objective function or standard for evaluating the quality of tree models. Since the main applications of the tree models are for entertainment purposes, not for precise duplication or detailed simulation, the visual inspection of the presence/absence of target features can be used to evaluate the quality of results and the effectiveness of algorithms. Actually the method of visual inspection and comparison has been widely used by many researchers (Stava et al., 1998; Runions, 2008; Xu and Mould, 2012). Here it is used to evaluate the results in this paper. Figure 10 shows a 3D tree with feature enhancement (on the left side) and a tree created using the previous method (Xu and Mould, 2015) without feature enhancement (on the right). The top row shows the tree skeleton

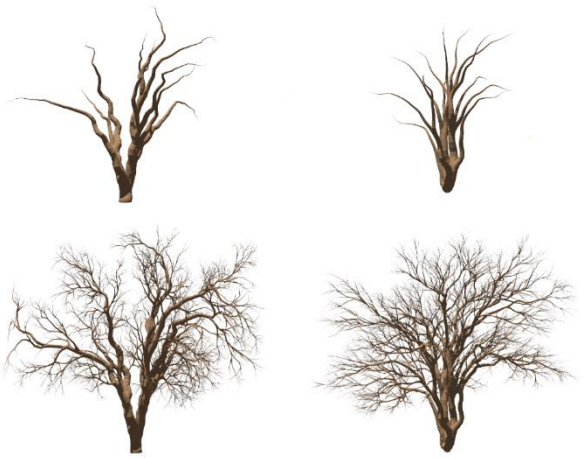


Fig. 10. Left: a tree with feature enhancement; right: a tree without feature enhancement from the early work of guiding vectors (Xu and Mould, 2015).

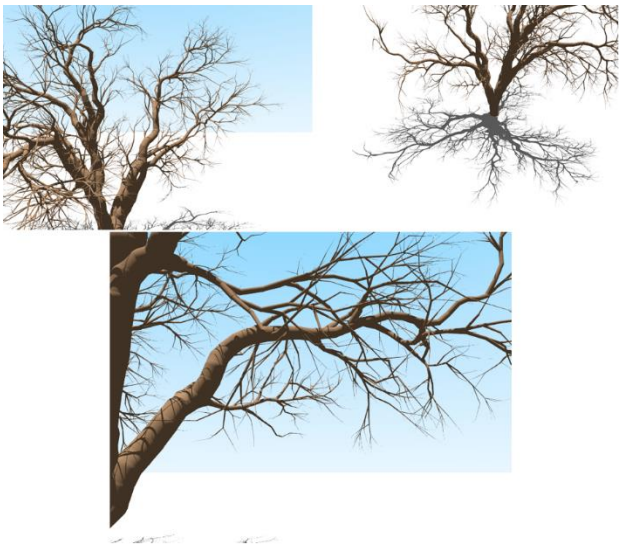


Fig. 11. A 3D tree with enhanced features viewed from three different directions.

and the lower row shows the trees after adding twigs. The left tree has more crooked branches which enhance the natural irregularities. The big gap in the central region of the left tree also makes the whole tree structure different from the right regular structure. This feature is hard for the previous method of guiding vectors. While the right tree shows more homogeneous properties, the left tree possesses more heterogeneous features. Figure 11 shows a 3D tree viewed from three different directions. A close view of individual branches is in the lower image, where due to the feature enhancement, the branch demonstrates a crooked shape that spreads from the lower left to the upper right and curves upwards.

5 Conclusion and Future Work

This paper presents a feature enhancement algorithm for 3D tree synthesis. The algorithm is built on the idea of path planning and tree synthesis with guiding vectors. In this algorithm, a number of holes are placed in the graph, and edges in the holes are disconnected thus prevent the least-cost paths to cross the regions of holes. Due to the limited available edges in narrow non-hole regions, the least-cost paths are forced to take edges in non-hole regions especially those are close to the contour of holes, leading to crooked path shapes. The hole size contributes to control the scale of this irregularity: large holes make branch gaps and small holes make the fine crookedness.

Compared with the previous work, the algorithm can effectively enhance the branch features and make branch shapes more irregular. The use of holes, including the hole distribution and hole sizing, contribute to enhance the heterogeneity of branches especially the natural gaps between branch clusters. These features are helpful to create irregular 3D trees which are needed in movie scenes or game animations especially for special scenes where irregular crooked trees (such as Halloween spooky trees) are required. The algorithm also has its limitations. The current combination of large holes and small holes is tentative. A more sophisticated and systematic investigation of the proportion should be performed in the future work. In addition, the distribution of holes is based on Poisson Disc distribution, which however does not take the overall tree structure into consideration. Different regions of the tree structure can be treated differently, with different densities of holes. In the future work, more interactive controls, such as sketch-based methods, will be explored to make the feature enhancement more controllable.

Acknowledgements

Thanks Dr. David Mould for valuable discussions about graph modification ideas. This work has been supported by the Organized Research and Creative Activities (ORCA) Program of University of Houston-Downtown.

Conflict of Interest: none declared.

References

A. Runions. Modeling biological patterns using the space colonization algorithm, 2008

D. S.Ebert, F. Musgrave, D. Peachey, K. Perlin, and S. Worley. Texturing and Modeling, A Procedural Approach (Third Edition). Elsevier Science, USA, 2005.

E. Fiume. An Introduction to Scientific, Symbolic, and Graphical Computation. A K Peters/CRC Press, 1995.

E. W. Dijkstra. A note on two problems in connexion with graphs. Numerische Mathematik, 1:269–271, 1959.

H. Xu, N. Gossett, and B. Chen. Knowledge and heuristic-based modeling of laser-scanned trees. ACM Trans. Graph., 26, 2007.

I. Millington and J. Funge. Artificial Intelligence for Games, Second Edition. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition, 2009.

L. Xu and D. Mould. Modeling dendritic shapes - using path planning. In GRAPP (GM/R), pages 29–36, 2007.

L. Xu and D. Mould. Synthetic tree models from iterated discrete graphs. In Proceedings of Graphics Interface, pages 149–156, 2012.

L. Xu and D. Mould. Procedural tree modeling with guiding vectors. Computer Graphics Forum, 34(7), 2015.

O. Stava, S. Pirk, J. Kratt, B. Chen, R. Mch, O. Deussen, and B. Benes. In search of the right abstraction: The synergy between art, science, and information technology in the modeling of natural phenomena. Art @ Science, pages 60–68, 1998.

P. Prusinkiewicz and J. Hanan. Lindenmayer systems, fractals and plants. Lecture Notes on Biomathematics, 1989.

P. Prusinkiewicz and A. Lindenmayer. The Algorithmic Beauty of Plants. SpringerVerlag, New York, 1990.

T. Capizzi. 3D Modeling And Texture Mapping. Premier Press, 1982.

Y. Li, X. Fan, N. J. Mitra, D. Chamovitz, D. Cohen-Or, and B. Chen. Analyzing growing plants from 4d point cloud data. ACM Trans. Graph., 32(6):157:1–157:10, 2013.

Y. Livny, F. Yan, M. Olson, B. Chen, H. Zhang, and J. El-Sana. Automatic reconstruction of tree skeletal structures from point clouds. ACM Trans. Graph., 29(6):151:1– 151:8, 2010.

American Journal of Advanced Research, 2017, 1–1
Sept. 2017, pages 7-12
doi: 10.5281/zenodo.836235
http://www.ibii-us.org/Journals/AJAR/
ISBN 2572-9829 (Online), 2572-9810 (Print)

Rain Attenuation Analysis from System Operating at Ku and Ka Frequencies Bands

Osahenvemwen, O.A*, and Omorogiuwa, O
Department of Electrical and Electronic,Faculty of Engineering and Technology
Ambrose Alli University, Ekpoma, Nigeria.

*Email: Osahenvemwenaustin@ymail.com

Received on May 12, 2017; revised on July 02, 2017; published on July 29, 2017

Abstract
This study presents rain attenuation analysis from System operating at Ku and Ka Frequencies Bands. Focused to determine the effects of weather conditions on propagation link performance property and quality of service render to customers. Recently, there is high demand of data rate, voice and video information have being on the increase due to industrial evolution and technological advancements. Satellite communication systems have witnessed various impairments which are considered in this study. Basic climate data were obtained from Meteorological Centre in Nigeria, for three years duration (from January 2013 to December 2015). Four geographical locations, Abuja, Benin City, Kano and Warri were considered in this study, in Nigeria. It is observed that Warri witnessed the highest rainfall followed by Benin City, Abuja and lastly by Kano. It is observed that the month of August witnessed the highest rainfall per each month followed by September and July at the year 2015. It was observed that rainfall distribution pattern possess a non–linear distribution pattern. Therefore, Empirical modelling was deployed; using excels software to determine the rainfall pattern. Based on the analysis, it was observed that rainfall followed exponential distribution pattern of rainfall. The basic parameters used in rain attenuation were highlighted with the corresponding attenuation characteristics such has rain intensity, rain height, effective path length, specific attenuation, rainy path length, slang path and horizontal projection specific attenuation were determined.

Keywords: Empirical modelling, Exponential, Rain Attenuation, and Rainfall Distribution Pattern

1 Introduction

The uses of data, voice, video information has being on the increase due to industrial evolution and technological advancements. The demand for multimedia applications, advancement in broadcast industries and quest for high quality of service performance has become imperative in this past decade. Also, with rising competition among internet and communication (voice) service providers to woo more subscribers with different applications and high quality of service render. Lastly, in view of the difficult terrain of some subscriber's locations, become necessary to use satellite technology to provide reliable quality of service. Radio wave propagation between terrestrial and earth space links are adversely affected by atmospheric gases, rain, and cloud, fog and atmospheric scintillation. The problem becomes more acute for system operating at frequencies above 10GHZ (Ku, Ka and V bands). Nigeria is located in the tropics unlike the temperate environments such as Europe and North America. The effects of the troposphere in micro wave signals will be most severe in the tropics because of high occurrence of rainfall, high temperature, high relative humidity and high rain intensities which lead to a more severe signal attenuation and outage then what is experienced in the temperate region. Unfortunately, most communication satellite equipment supplied to Africa are designed based on propagation measurement made in Europe and North America, which may not meet the minimum availability standard that can cope with hash weather experienced in the tropics.

The difficult terrains such as swamp areas, riverine areas are associated with heavy rainfall. The C-band (6/4 GHz) satellites communication were deployed, but due to some drawbacks associated, such as bandwidth limitation (within 500 MHz - 1000 MHz), interference due to congestion, narrow beam cannot be transmitted and direct reception in TVs at home uses bigger sizes of parabolic dishes (Sapna, 2008). Which necessitate, the use of Ku band (under k- band 14/12 GHz), Ka band (after k-band 30/20 GHz) and V band (70/55 GHz) take the centre stage of satellite communication due to higher performance, relatively congestion free. The satellite communications using frequencies of higher than 10 GHz (such has Ku band, Ka band and V band) are affected by various propagation impairments such as rain attenuation, delay, cloud attenuation, Tropospheric scintillation, Ionospheric scintillation, water vapor attenuation, and rain and ice depolarization (Sridhar, *et al* 2012, Osahenvemwen, 2013).

Rainfall or raindrop results in attenuation of radio waves by scattering and by absorption of energy from the wave and increase with increase in frequency, which degrades the reliability and performance of the communication link. (Sapna, 2008). Rain effects are dependent on frequency, rain rate, rain drop size distribution and drop shape, which are determined by the type of rain witnessed in this region under investigated (Yee Hui Lee_ and Stefan Winkler).The country Nigeria lie in tropical heavy rain regions with coordinate of latitude 10^0 00'N longitude $8^0$00'E with total area of 923,768 square kilometers. It located within the equator

and tropic of cancer the latitude of Nigeria fall within the tropical zone but the climate conditions are not entirely tropical in nature. The climate condition varies in most parts of the country. In the north the climate condition is arid and to the south there is an equatorial type of climate. Therefore, due to this variation in climate condition on different geographical locations in Nigeria, it is important to carry out holistic study on this environment.

There are two common rain types, which are convective rain and stratiform rain. The Convective rain arises from vertical atmospheric motions resulting in vertical transport and mixing. The convective rain flow occurs in a cell whose horizontal extent is usually several kilometers. The cell may be isolated or embedded in a thunderstorm region associated with a passing weather front. Because of the motion of the front and the sliding motion of the cell along the front, the high rain rate usually only lasts for several minutes. These rains are the most common sources of high rain rate events. While, the stratiform rain typically shows a stratified horizontal extent of hundreds of kilometers, durations exceeding one hour, and low rain rates of less than 25 mm/hr. For communication applications, stratiform rain occurs for sufficiently long periods of time that a link margin may be required to exceed the resulting attenuation. Stratiform rain covers large geographic areas, and the spatial distribution of total rainfall is expected to be uniform. During stratiform type of rain the point, where the frozen water or ice begins to melt and forms rain, is the 0^0c isotherm height (Ojo *et a*l). Likewise the rain rate averaged over several hours is expected to be rather similar for ground sites located up to tens of kilometers apart (Yee Hui Lee_ and Stefan Winkler). The two major causes of rain fade are:

Absorption – water molecules in a rain droplet absorb portions or all of the signal energy of the passing radio wave. With shorter wavelength, there will be more interaction between the radio wave and water molecules, leading to increased in energy losses.

Scattering – this is a physical process, caused by either refraction or diffraction, in which the direction of the radio wave deviates from its original path as it passes through a medium containing raindrops. This disperses the energy of the signal from its initial travel direction. The accumulation of these different reactions ultimately leads to a decrease in the level of received signal, thus resulting in rain attenuation (Osahenvemwen, 2013). In addition, rain Attenuation is being affected by variation of frequency, location, polarization and rainfall rate.

2 Satellite Propagation Technique

The problems observed on satellite transmission are due to absorption from the cloud, ice, rain attenuation due to snow, hail and fog. Satellite radio waves are propagated through ionosphere, troposphere and stratosphere. The ionosphere is the upper part of the atmosphere where sufficient ionization takes place, to influence radio wave propagation. The ionosphere usually consist three layers D, E and F region. D region is the lowest of the ionosphere layers, it extends from approximately 50 to 90km with the maximum electron of about $10^9/M^3$ occurring between 75 and 80km in the day time. As the electron concentration in the D region is very low, it tends to have little effect on light frequency wave. E region extends from 90 to140km and peak electron concentration occurs between 100 and 110km (Sanjeev, 2011).

Electron densities in E**0**region vary with the 11years sunspot cycle. Electron concentrations drop by a factor of about 100 at night in the E region. F region has the highest electron densities of the normal ionosphere. In the day time it normally consist of two parts F_1 and F_2. F_1 is largely disappear at night but has a peak density of about $1.5\times 10^{11}/M^3$ at noon the minimum of the solar cycle F_2 has the highest peak electron densities of ionosphere and it remains higher than in D and E region. Reflection from F_2 layer is the major factor in higher frequency communication which formally handled a large fraction of long distance especially transoceanic, communication (Muhammad.*et al*.2011).Since the ionization is mainly caused by solar radiation, it is dependent on the location, time of the day, season and sunspot. Radio wave's propagation through ionosphere experienced different attenuation at frequencies above 10GHZ. As reported in literature, atmosphere contains force electrons, ions and molecules which interact with radio waves dependent strongly on frequency, so as the frequency increases, the effect of attenuation increases.

The propagation factors that affect Ku and Ka bands satellite links communication includes, rain attenuation, depolarization due to rain and ice, antenna wetting, gaseous absorption cloud attenuation, wetting layers attenuation and tropospheric scintillation. In several rains fade compensation approaches can be considered at three major characteristics of the propagation factor as followed: rang of signal fading (dB), rate of signal fluctuation (dB/sec) and frequency dependence (scaling)

3 Methods

This study was on rain attenuation analysis from system operating at Ku and Ka frequencies bands. Using Nigeria as a case study, which lie in tropical heavy rain regions with coordinate of latitude 10^0 00' N and longitude 8^0 00' E with total area of 923,768 square kilometers. Its located within the equator and tropic of cancer the latitude of Nigeria fall within the tropical zone but the climate conditions are not entirely tropical in nature. Four geographical locations were considered within the country such are Abuja, Benin City, Kano and Warri, were considered in Nigeria. The daily rainfall data were obtained from Meteorological Center in Nigeria, for period of three year (from January 2013 to December 2015) for four different geography locations in Nigeria. Rainfall distribution pattern was determine, Empirical modeling and comparison between linear and exponential distribution pattern of rainfall for Benin City were considered in this research. Analysis of rain attenuation characteristics such has rain intensity, Effective path length, specific attenuation, rainy path length, rain height, slang path, horizontal projection specific attenuation were determined based on the ITU-R prediction model for various region based on obtained data and 12 GHz as operating frequency.

Fig1: Map of Nigeria shown the four different locations under investigation

3.1 Data presentation

The daily rainfall for three years for four different geography locations in Nigeria such are Abuja, Benin City, Kano and Warri, were considered

Table 1: Average Rainfalls mm (inches) for Three Year, Four Different Locations

Months	Abuja			Kano			Warri			Benin city		
	2013	2014	2015	2013	2014	2015	2013	2014	2015	2013	2014	2015
Janu	**0.1**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**3.5**	**2.9**	**0.0**	**0.4**	**2.7**	**0.9**
Feb	**1.4**	**0.1**	**0.0**	**0.0**	**0.0**	**0.0**	**0.6**	**0.6**	**3.3**	**7.2**	**1.9**	**6.2**
Mar	**0.9**	**2.2**	**2.3**	**0.0**	**0.0**	**0.0**	**4.9**	**4.7**	**4.5**	**4.1**	**4.2**	**5.9**
April	**4.5**	**5.2**	**0.8**	**1.5**	**0.6**	**0.0**	**6.9**	**5.9**	**7.9**	**6.7**	**7.1**	**3.6**

May	**4.0**	**7.7**	**7.7**	**1.3**	**2.4**	**0.0**	**7.5**	**11.5**	**10.8**	**10.1**	**8.1**	**6.7**
June	**5.4**	**5.8**	**5.8**	**2.6**	**3.0**	**2.3**	**15.3**	**10.2**	**16.0**	**8.5**	**8.1**	**13.3**
July	**6.2**	**2.4**	**2.4**	**5.1**	**15.5**	**4.9**	**16.0**	**6.3**	**6.1**	**12.6**	**8.8**	**12.5**
Aug	**5.9**	**17.9**	**4.5**	**13.3**	**16.4**	**15.6**	**5.9**	**11.3**	**20.2**	**5.4**	**13.1**	**7.1**
Sept	**9.1**	**7.0**	**4.4**	**5.9**	**6.2**	**3.9**	**6.2**	**13.7**	**11.1**	**18.8**	**12.3**	**11.4**
Octo	**6.5**	**7.0**	**7.0**	**0.3**	**0.6**	**0.3**	**4.9**	**15.3**	**16.9**	**10.9**	**11.8**	**8.8**
Nov	**0.0**	**0.1**	**0.0**	**0.0**	**0.0**	**0.0**	**3.8**	**8.4**	**7.6**	**3.5**	**5.4**	**0.6**
Dec	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.1**	**0.7**	**0.0**	**1.9**	**1.1**	**0.4**

Table 2 Climatological parameters of the stations under investigation

S/N	Locations	Latitude	Longitude	Height above sea level (Meters)	Elevation angle
1	Abuja	9^0 12' N	7^0 11' E	476 meters (0.476 KM or 1561 ft)	60.0^0
2	Benin City	9^0 30' N	2^0 15' E	86 meters (0.086 KM or 282ft)	60.0^0
3	Kano	12^0 02N	$08^0$30E	488 Meters (0.488 km) or 1,601 ft	30.0^0
4	Warri	5^0. 52` N	$5^{0.}$ 75` E	21 meters (0.021 KM or 69 ft)	60.0^0

Source: [http//dateandtime.info/citycoordinate php, 2015].

(Note) Elevation above sea level: 484 m = 1587 ft

3.2 Basic rain characteristics that affect satellite communication

Rain attenuation is defined as the product of specific attenuation (dB/km) and the effective propagation path length (km). The product of path reduction factor and the physical path length of a microwave link are referred to as the effective path length. The strength of satellite signal may be degraded or reduced under rain conditions; in particular radio waves above 10 GHz are subject to attenuation by molecular absorption and rain. Presence of rain drops can severely degrade the reliability and performance of communication links. Attenuation due to rain effect is a function of various parameters including elevation angle, carrier frequency, height of earth station, latitude of earth station and rain fall rate. The primary parameters, however, are drop-size distribution and the number of drops that are present in the volume shared by the wave with the rain. It is important to note that, attenuation is determined not by how much rain has fallen but the rate at which it is falling. The propagation loss due to rain is given by:

$$L = 10\log\frac{p_r(0)}{p_r(r)} \quad (1)$$

Where P_0 is the signal power before the rain region, Pr is the signal power after the rain region, and r is the path length through the rain region. The propagation loss due to rain attenuation is usually expressed by specific attenuation γ in decibels per kilometer, so propagation loss is:

$$L = \gamma l_r \quad (2)$$

Where γ is specific attenuation in dB/km and *lr* is rain path length in km. Based on ITU-R specific attenuation model. It is found that γ depends only on rainfall rate, measured in millimeters per hour. From this model, the usual form of expressing γ is:

$$\gamma = KB^{\alpha}[dB/KM] \quad (3)$$

Where K and α is frequency dependent coefficients.

Attenuation can be obtained from direct measurements or predicted from the knowledge of rain rate.

In determining rain attenuation along the slant-path of satellite communication required the rain parameters as followed:

F: the frequency of operation in GHz, (12 GHz)

θ: the elevation angle to the satellite, in degrees, (in Table 2)

Ø: the latitude of the ground station, in degrees N and S, (in Table 2)

hs: the height of the ground station above sea level, in km, (in Table 2)

Re: effective radius of the Earth (8 500 km),

$R_{0.01}$: point rainfall rate for the location of interest for 0.01% of an average year (mm/hr).

Firstly to compute the rain attenuation along the slant-path in satellite communication system is summarized as follows:

Step 1: Determine the rain height, h_R (KM) from the recommendation ITU-R P.839 as

$$h_R = ho + 0.36KM \quad (4)$$

Where ho is the 0° C isotherm height above mean sea level at the desired location

Step 2 Determine the slant-path length (L_S) km

$$L_s = \frac{h_r - h_s}{\sin\theta} \text{ for } \theta \geq 5^0 \quad (5)$$

Also, it given as

$$L_s = \frac{2(h_r - h_s)}{(sin^2\theta + \frac{2(h_r-h_s)}{R_e})^{\frac{1}{2}}} \quad \text{for } \theta \leq 5^0 \quad (6)$$

Where θ is Elevation angle in degrees, h_s is the height of the location above sea level in Km, and h_R is the rain height in Km.

The horizontal projection is then expressed as:

$$L_G = L_S\cos\theta \quad (7)$$

Step 3: Determine the rain rate at 0.01% for the location of interest (Warri) over an average year. In this work, 62mm/hr from is used to derive rain rate at one-minute integration time at 0.01% of exceeded from long-term local data.

Step 4: Calculate the specific attenuation, γ_R, by using the frequency dependent regression coefficients provided in

ITU-R P.838 Recommendation $R_{0.01}$ and using the following Equation

$$\gamma_R = K(R_{0.01})^{\alpha} \quad \text{dB/km} \quad (8)$$

K and α are regression coefficients which depend on the frequency, polarization, raindrop size distribution and temperature and obtained using,

$$K = \frac{K_H K_v + (K_H - K_v)cos^2\theta\cos(2t)}{2} \quad (9)$$

$$\alpha = \frac{K_H\alpha_H + K_v\alpha_v + (K_H\alpha_H - K_v\alpha_v)cos^2\theta\cos(2t)}{2K} \quad (10)$$

Where t is the polarization tilt angle relative to horizontal and θ is the path elevation angle. The polarization tilt angle t = 90 for vertical polarization and t=0 for horizontal polarization while circular polarization is given as t = 45. The frequency dependent coefficients k_H, k_v, α_H and α_v are obtained for both horizontal and vertical polarization over frequencies of 1-400 GHz.

Step 5: Determine the horizontal reduction factor or horizontal path adjustment factor, $r_{0.01}$ at 0.01% probability and expressed as:

$$r_{0.01} = \frac{1}{1 + 0.78\sqrt{\frac{L_G\gamma_R}{f}} - 0.38(1 - e^{-2LG})} \quad (11)$$

Note: L_G is the horizontal projection as determined in Step 2 and *f* is the operating frequency measured in GHz.

***Step 6*:** calculate the adjusted rainy path length L_R (KM) through rain using

$$L_R = \frac{L_G r_{0.01}}{COS\theta} \quad \text{for } \xi > 0 \quad (12)$$

$$L_S = \frac{h_R - h_s}{sin\theta} \quad \text{for } \xi \leq 0 \quad (13)$$

$$\xi = \tan^{-1}\frac{h_R - h_s}{L_G r_{0.01}} \text{ Degrees} \quad (14)$$

Step 7 obtain the vertical reduction factor or vertical adjustment factor, $v_{0.01}$, for 0.01% of the time by using

$$V_{0.01} = \frac{1}{1 + \sqrt{sin\theta}\left[31\left(1 - e^{-\frac{\theta}{1+X}}\right)\frac{\sqrt{L_R\gamma_R}}{f^2} - 0.45\right]} \quad (15)$$

Where X=36-/ Ø/,for /Ø/<36^0 (16)

X=0, for / Ø/> 36^0 (17)

***Step 8*:** The effective path length through rain L_E is then computed from:

$L_E = L_R V_{O.O1}$ KM (18)

***Step* 9**: Calculate the predicted attenuation exceeded for 0.01% of an average year.

$A_{0.01} = \gamma_R L_E$ DB (19)

***Step* 10:** The estimated attenuation to be exceeded for the other percentages of an average year, in the range 0.001% to 10% may then be estimated using $A_{0.01}$ as

$A_P = A_{0.01}(\frac{P}{0.01})^{-[0.655+0.03In(P)-0.045In(A_{0.01})-\beta sin\theta(1-p)]}$ (20)

Where p is the percentage probability of interest and β is given by

For $P \geq 1\%$, $\beta = 0$ (21)

For $p < 1\%$, $\beta = 0$ if $/Ø/ \geq 36^0$ (22)

$\beta = -0.005 (/Ø/-36)$ for $\theta\ 25^0$ and $/Ø/ \leq 36^0$ (23)

$\beta = -0.005(/ /-36) + 1.8-4.25\sin\theta$ (24)

For $\theta < 25^0$ and $/Ø/< 36^0$ (25)

Table 3. Various rainfall rates with different % time exceeded (Warri)

S/N	% Time exceeded	Rainfall Rate (mm/hr)	Attenuation (dB)
1	0.001%	114.36	35.45
2	0'01%	62.00	16.35
3	0.1%	17.13	3.22
4	1%	1.94	0.21
5	10%	0	0.0

In this analysis some parameter were assumed to be able to determine the rain attenuation parameters, such operating frequency at 12 GHz, and other basic parameters obtained are from Table 2.

Table 4. Rainfall Parameters Related to Ku and Ka frequencies bands.

S/N	Locations under investigation	Duration of years	Rainfall (mm)	rain height	Effective path length, (Km)
1	Abuja	2013-2015	9.25	4.17 km	12.03
2	Benin City	2013-2015	12.36	3.78 km	10.23
3	Kano	2013-2015	8.35	4.18 km	12.11
5	Warri	2013-2015	13.14	3.71 km	10.07

Determine the slant-path length (L_S) km or effective path length

$L_s = \frac{h_r - h_s}{\sin\theta}$ for $\theta \geq 5^0$

$L_S = \frac{3.78-0.086}{Sin\ 60^0}$ =10.23km For Benin City location

The slant-path length (L_S) km or effective path length is used to determine the rain attenuation in satellite communication system.

4 Data Analysis, Result and Discussion

The data obtained from Nigeria metrological centre from 2013 to 2015 for four different geographical locations are analysis using Microsoft excel software to determine various Figures.

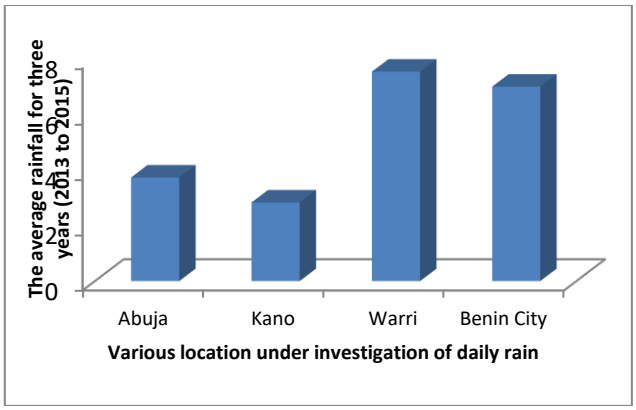


Fig 2: total rain fall for three years for four different geographical locations

The daily rainfalls were obtained from metrological centre in Nigeria, from the three years data, from 2013 to 2015 were considered for four geographical locations. Warri and Benin City are from the south of Nigeria, while Abuja was from the centre part of Nigeria and lastly Kano was located in north part of Nigeria. It was observed that warri witnessed the highest rainfall has shown in Fig 2, followed by Benin City, Abuja and lastly by Kano.

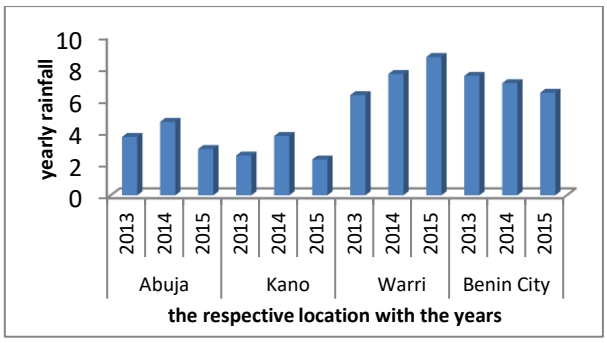


Fig 3 cascaded of rainfall into yearly bases on the four different geographical locations

The rainfalls were cascaded into yearly bases in the four different geographical locations in Nigeria. From the Fig 3, the year rainfalls were compared; it was observed that Benin City, Kano and Abuja witnessed lower rainfall throughout the 2015 year, while Warri witnessed the heaviest rainfall in 2015. In Fig 4 showed the rainfalls in four different locations in relative to years and months. It was observed that the month of August witnessed the highest rainfall per each month followed by September and July

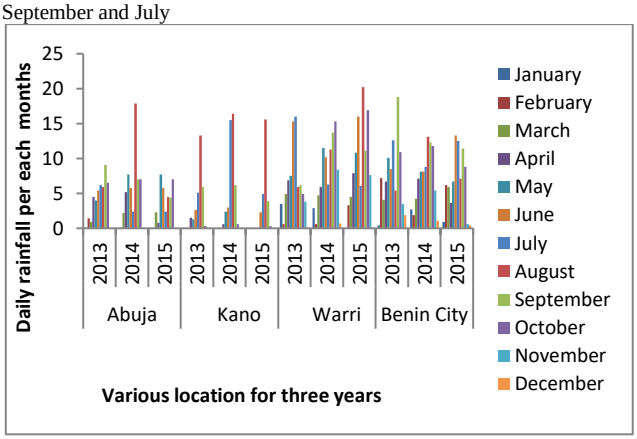
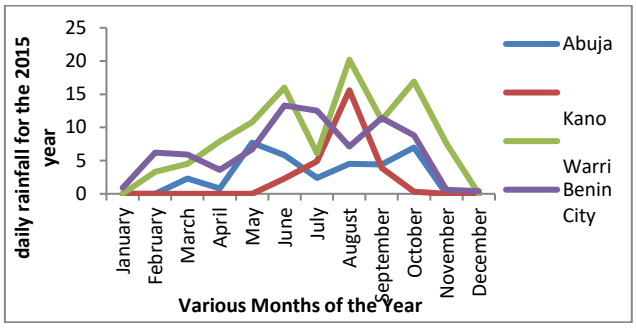


Fig 4 Rainfalls in Four Different Locations with respectively Years and Months



In Fig 5 presents the rainfall distribution pattern for four different geographical locations, the rainfall were capture throughout the days and the months in 2015 years and was compared with rainfall distribution pattern of Benin City location, using different years presented in Fig 6. It was observed that rainfall distribution pattern is non–linear distribution pattern.

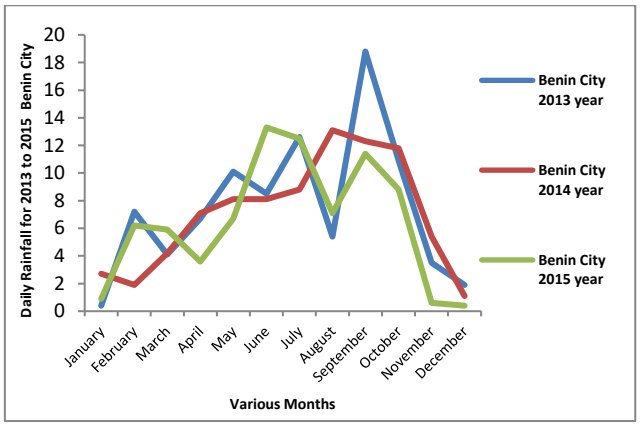


Fig 5 the rainfall Distribution pattern for Benin City for three years

4.1 Empirical modeling

Empirical modeling refers to any kind of computer modeling based on empirical observations rather than on mathematically describable relationships of the system modeled. Based on the data obtained, and developing empirical model, deploying micro software using excel package, the resultant output were presented in Fig 7 and Fig 8. This empirical model was developed based on rainfall data obtained from Benin City location for three years. The distribution pattern was examined as Exponential distribution pattern and the resultant equations were presented in Fig 7, similarly linear distribution pattern was considered in Fig 8. The comparison of between Exponential distribution pattern and linear distribution pattern are presented in Table 4.

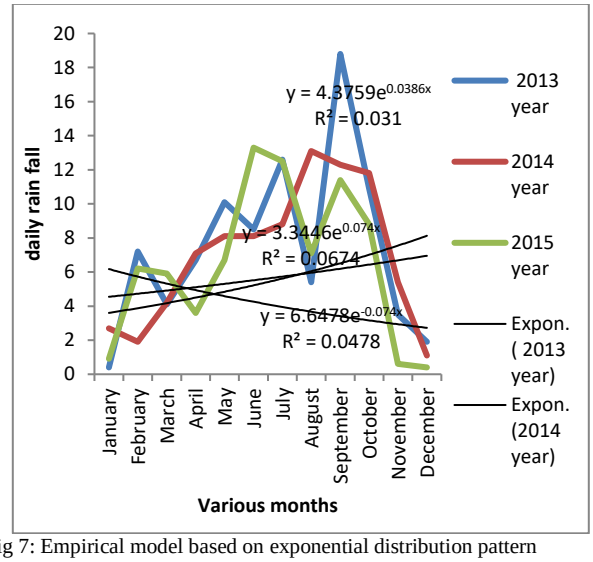


Fig 7: Empirical model based on exponential distribution pattern

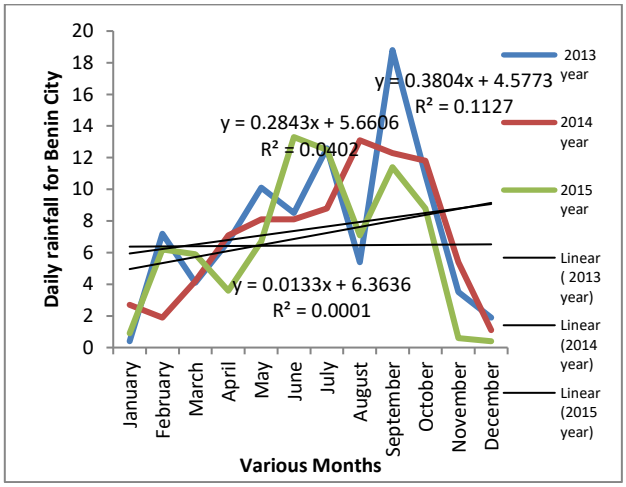


Fig 8 Empirical model based on linear distribution pattern

Table 5: the comparison between linear and exponential distribution pattern of rainfall for Benin City

S/N	Years	Linear Distribution Pattern		Exponential Distribution Pattern	
		Resultant Equation	R^2-Squared value	Resultant Equation	R^2-Squared value
1	2013	$y = 0.28x + 5660$	$R^2 = 0.040$	$Y = 3.344e^{0.07x}$	$R^2 = 0.067$
2	2014	$y = 0.389x + 4.577$	$R^2 = 0.112$	$y = 4.375e^{0.038}$	$R^2 = 0.031$
3	2015	$y = 0.013x + 8.36$	$R^2 = 0.000$	$y = 6.647e^{-0.07x}$	$R^3 = 0.047$

The resultant outputs from empirical modeling are presented in Table 5, which involves the comparison between linear and exponential distribution pattern of rainfall for Benin City for duration of three years (2013-2015) average rainfall. It was observed that exponential distribution

pattern of rainfall were more consisted, compared to the result obtained from linear distribution rain pattern in Table 5.

Therefore the exponential distribution pattern Equation is presented in Equation (26).

$$y = 4.375e^{0.038x} \quad (26)$$

Y represents the average rainfall from 2013 to 2015; x represents the various months of the year. The resultant R^2- squared value obtained is 0.031.

4.2 Analysis of rain attenuation parameters in respect to satellite communication

Nigeria is a tropical region that lies between the geographical area of 4°N, 3°E and 14°N, 15°E in the South-Eastern edge of the West African region. It is characterized by rainy and dry

4.2.1 Analysis of rain attenuation parameters in respect to satellite communication

Nigeria is a tropical region that lies between the geographical area of 4°N, 3°E and 14°N, 15°E in the South-Eastern edge of the West African region. It is characterized by rainy and dry seasons. The rainy season (between the months of April and October) is heavily influenced by an air mass originating from the South Atlantic Ocean, known as the South-West (monsoon) wind or the Tropical Maritime (MT) air mass, while the dry season (between November and March) are accompanied by a dust laden air mass from the Sahara desert, known as Harmat-tan or the Tropical Continental (CT) air mass. Harmattans (between December and February) are the dry, hot-by-day and cold-by-night dust-laden North-East wind, which is usually associated with the variable intensification of the sub-tropical anticyclone. Nigeria, like any other tropical region is characterized mostly by convective rainfall types (Convective, Monsoon Precipitation and Tropical Storm).

5 Conclusion

Most satellite services are offered at C band frequencies, but a lot of interest applications emerged in deploying Ku and ka-band satellites. The advantages offered, by ku and ka band satellite services include terrain independence and large coverage areas etc. The congestion experienced at lower frequency bands satellite, due to high demand for diverse services and increased bandwidth used by voice calls were explored by higher frequency bands (ku and ka band). The major drawback experienced in operating at such a higher frequencies (Ka-band and above), is signal distortion result from free space and rainfall. In this study, the rain attenuation is calculated, for different rainfall rates and exceedence percentages under the investigated area. It is observed that attenuation increases as rainfall rates increase. However, rainfall attenuation is strongly dependent on two factors as followed: operating frequency and the local rain rate. Also, the effective path length for each region is determined and it is dependent on frequency and elevation angle. The effective path length depends strongly on elevation angle and operating frequency of the satellite. Also the vertical polarization produces less rain attenuation than horizontal polarization at Ku and Ka band and the rainfall distribution pattern followed exponential distribution pattern under area investigated. Further study, can be carried out in this area by developing models for rain attenuation parameters in corresponding to the locations under investigation.

References

Badron, K. Ismail, A.F. Islam, M.R. Tharek, A.R and Din, J. (2011). Rain Attenuation Prediction Model for Tropical V-band Satellite Earth link, International Conference on Telecommunication Technology and Applications Proc .of CSIT, Vol.5 IACSIT Press, Singapore.

http//www.brifannica.comn/EBchecked/topic/524891/satellite-communication,Satellite communication, accessed on 4th July, 2013.

http//en.wikipedia.org/wiki/communications- satellite, communication satellite, accessed on 6th June, 2013.

http//www.brifannica.comn/EBchecked/topic/524891/satellite-communication, Satellite communication, accessed on 4th July, 2013.

http//en.wikipedia.org/wiki/warri, Warri, accessed on 7th July, 2013.

http://jwcn.eurasipjournals.com/content/2012/1/189, Islam et al. EURASIP Journal on Wireless Communications and Networking 2012, 2012:189, accessed 26 /6/2013.

http//www.nairalandcom/7715/official-free-to-air, Office thread of free to air satellite TV (part 1), accessed on 21st July, 2013.

http//www.nimetring.org/uploads/publication/Annual,2010 climate review Mkparu, accessed on 6th July,2013.

http// www.map- streeviewcomNigeria/warri;Warri-map, accessed on 9th July, 2013

http//www.nimet.org/uploads/publication/nimet, Seasonal rainfall prediction (SRP), accessed on 2nd August, 2013.

Kirtay, S. (2002). Broadband satellite system technologies for effective use of the 12-30GHz radio spectrum, Electronics and Communication Engineering Journal page 79-89.

Isikwue, B. C., Ikoyo, A.H. and Utah, E.U, (2013). Analysis of rainfall rates and attenuations for Line–of–sight EHF/SHF radio communication links over Makurdi, Nigeria, Research Journal of Earth and Planetary Sciences Vol. 3(2) pp. 60 – 74.

Malinga, P. A. Owolawi, and T. J. O. Afullo, (2013). Estimation of Rain Attenuation at C, Ka, Ku and V Bands for Satellite Links in South Africa, PIERS Proceedings, Taipei,28,

Muhammad Zubair, Zaffar Haider, Shahid, A. Khan, Jamal Nasir, (2011). Atmospheric influences on satellite communications COMSATS Institute of Information Technology, Islamabad (1,2,3), COMSATS Institute of Information Technology, Abbott Abad (4), Przegląd Elektrotechniczny (Electrical Review), ISSN 0033-2097, R. 87 NR 5.

Ojo, J. S. and Ajewole, M. O. (2008). Rain rate and Rain attenuation prediction for satellite communication in Ku and Ka bands over Nigeria, Progress in Electromagnetics Research B, Vol. 5, 207–223.

Osahenvemwen, O.A. and Omorogiuwa, O. (**2013),** Effect of Rain on Satellite Communication Networks in Nigeria: Case Study of Warri Town, Journal of Nigeria Association of Mathematical Physics, Volume 25,No2**,** Pp107-114. (http://nampjournals.org/abstracts/vol25abstract.pdf)

Sanjeev, S. J., Suresh and Ganesh Madhan, M., (2011). An Experimental Study of Rain Attenuation at Ku Band Frequencies for an Earth Space Path over Chennai, International Symposium on Devices MEMS, Intelligent Systems & Communication (ISDMISC) Proceedings published by International Journal of Computer Applications (IJCA), page 46-49.

Sridhar, M. K., Padma Raju, and Ch. Srinivasa Rao, (2012). Estimation of Rain Attenuation based on ITU-R Model in Guntur (A.P), India, ACEEE Int. J. on Communications, Vol. 03, No. 03,

Vatcharakorn, Netharn, Surasee Prahmkaew and Siriwhaddhanah Pongpadpinit, (2013). Improve Receive Signal over Ku-Band Satellite Communications Based on Fuzzy Logic, International Journal of Scientific & Engineering Research, Volume 4, Issue 1, ISSN 2229-5518,Pp 45-48.

Yussuff Abayomi, I.O, Hisham Haji Khamis, (2012). Rain Attenuation Modelling and Mitigation in the Tropics: Brief Review, International Journal of Electrical and Computer Engineering (IJECE), Vol.2, No.6, Pp. 748~757.

Yasuyuki Maekawa, Tadashi Fujiwara1, Yoshiaki Shibagaki,(2006). Effects of Tropical Rainfall to the Ku-Band Satellite Communications Links at the Equatorial Atmosphere Radar Observatory, Journal of the Meteorological Society of Japan, Vol. 84A, Pp. 211-225.

American Journal of Advanced Research, 2017, 1–1
Sept. 2017, pages 13-18
doi: 10.5281/zenodo.910819
http://www.ibii-us.org/Journals/AJAR/
ISBN 2572-8849 (Online), 2572-8830 (Print)

Design and Implementation of a Computer Based Test Centre Using Biometric for Authentication

O. Omorogiuwa[1, *] and F. N. Nwukor[2]

[1]Department of Electrical and Electronic Engineering, Ambrose Alli University, Ekpoma, Nigeria., [2]Department of Electrical and Electronic Engineering, Petroleum Training Institute, Effurun, Nigeria.

*Email: owenseme@yahoo.com

Received on May 12, 2017; revised on Sept 20,2017; published on Sept 21, 2017

Abstract

The aim of this study is to develop a web-based computer centre using biometric finger print for verification/authentication and tie the MAC adders of all the system to the program server, tie the computers with the sever through quantum mechanics distribution (QKD) for the intranet only, to prevent intruders via intranet; building a system that will not be compromised and with desired confidence. Visual Studio was utilized to develop the Front-end aspect of the Computer Based Test (CBT) authentication software while Visual Basic.NET was applied to develop the software. The Backend was designed using MySQL server application, thus all the data for the CBT authentication software are stored in the MySQL database. During processing, records are retrieved from the database and displayed in the Front end. The performance under test was found to be satisfactory when comparing Manual verification/authentication (average 6.6sec) to Biometric verification/authentication (average 1sec). All unauthorized users are blocked and appropriate warning messages sent to the client by the server when they initiate Login procedure. This eliminates external access from unauthorized persons sitting for the examinations. All systems can see the network Identification, but not all persons have access right to the examination.

Keywords: Biometric, Finger print, MAC (Media Access Control) Address, CBT (Computer Based Test), IP (Internet Protocol), Database

1 Introduction

Computer based centre (CBC) is the connection of computers to a designated server through a switch or an array of switches for conducting of computer based examinations. For a CBC to be operational the computers must be able to communicate with the server in real time, as such system security will be required to protect the integrity of the system. In a CBC system, we have the hardware and software. The hardware includes the computers, switch, cable connectors, servers etc. The software represents the program that runs on the systems, enabling the user and the system to communicate. There have been many security systems that are tied to user interface. CBC, for example has a one-way authentication/verification, two-way authentication/verification and an IP authentication/verification, but with this authentication/verification there is still intruder in the system. In this paper we will be using IP/MAC for surveillance and Biometric fingerprint for authentication/verification which will be tied to the CBT.

Examination is one of the best methods of evaluating the knowledge and ability of an individual (Adebayo and Abdulhamid, 2014). Its purpose is to assess how much each student has learned compared to fellow students in the same course or learning situation. Various examination methods are being used in higher education institutions to assess academic progress, such as paper-pencil-based examinations, assignments, presentations etc. With the increasing rate of examination malpractices in the educational sectors the school management deserve to inculcate a tight security means to ensure that these activities of examination impersonators are reduced. The activities of these examination impersonators have harmed the educational sector greatly (Luecht, 2005a; 2005b). With the recent advancement in Information and Communications Technology (ICT), there are some drawbacks that can affect computer based centres with hackers always looking for loopholes in how examinations are conducted. Several approaches (Todorov, 2007) for security have been adopted by various individuals and institutions, such as one-way authentication and two-way authentication systems, and IP address for authentication (Ahmad and Abu, 2013; Saliu *et al*., 2013; Nafiu, 2014). A major drawback with the use of IP address on a system/network is that, it can be changed by hackers and this can compromise the security of a network system there by encouraging malpractice (Beaver, 2013). Other well-known security methods are firewalls, pass-wording, mathematical algorithms of encryption and decryption scheme.

The goal of this paper is to develop a computer based centre using biometric finger print for verification/authentication and tie the MAC adders of all the system to the program server for the intranet/internet, to prevent intruders via intranet/internet while building a system that will not be compromised and with desired confidence. The proposed system will use finger print biometrics. This would help ensure that only registered student during registration with their finger print are allowed into the examination hall, start examination and submit examination. The system without registered MAC adders on the server cannot be used for the examination, a scenario where you are your own access key to your examination. As one enables his/her examination it registers as attendance signed and as you submit using biometric you also sign out the attendance, this could seem very convenient.

The developed system would contribute in preventing malpractices in the educational sector. Impersonation which has eaten the educational system, thereby encouraging laziness among students would be eliminated and standard of student educational performance would be increased. Up till now the JAMB UTME, does not encourage the use of fingerprint as mode of authentication, this has resulted in people sitting for examinations for others. With the adoption of fingerprint, irregularities will be eliminated as fingerprint authentication will also be employed during collection of results and certificates. Nevertheless, for some people, the use of fingerprint technology is very intrusive, because is still related to criminal identification. Moreover, there can be inherent errors in the system with the dryness (or dirty) of the finger's skin, as well as with the degradation due to aging. In addition, a fingerprint can be copied and used by hacker. Table 1 shows a comparison of other biometric options that can also be used. In general, most persons prefer to use finger print biometric than other biometric like iris and retinal scan because of the talk of the risk that it may affect the human eyes.

Table 1. Comparison of different biometric technology

Biometric technology	Accuracy	Cost	Devices required	Social acceptability
Iris recognition	High	High	Camera	Low
Retinal scan	High	High	Camera	Low
Finger print	High	medium	Scanner	Medium

2 Methods

The user is enrolled into the system using biometric and the enrolment is tied to the CBC program platform, this is stored as a template on the database. When a user attempts to enter the examinations platform, a biometric program will pop up for the user to put his finger on the scanner for verification/authentication and the main features of the object scanned are then extracted and converted into a digital representation. This file is then compared to the templates on the database. If a match is found, the user is granted access to the examination platform.

For this to be achieved, the front-end aspect of the CBT authentication software is designed with Visual Studio. The language used to develop the software is Visual Basic.NET. The backend is designed with MySQL server application, which is easy-to-use database software. All the data for the CBT authentication software are stored in the MySQL database. Then records are retrieved from the database and displayed in the front end. A connector called mysql connector is installed in the computer system to provide a way for the visual basic.net codes to be able to access the data stored in the database. The connector provides access to the tables and records and they are copied to the computer memory.

2.1 Finger Print Technique (http://www.androidauthority.com/how-fingerprint-scanners-work-670934/)

The minutiae extraction technique implemented is based on the widely employed Crossing Number method. For the image post-processing stage I implemented the minutiae validation algorithm. The fingerprint scanner that was use is capacitive scanners**,** as capacitors can store electrical charge, connecting them up to conductive plates on the surface of the scanner allows them to be used to track the details of a fingerprint. The charge stored in the capacitor was changed slightly when a finger's ridge is placed over the conductive plates. An op-amp integrator circuit was used to track these changes, which was then recorded by an analogue-to-digital converter. Figures 1 shows the circuit diagram of the theory and architecture behind a capacitive fingerprint scanning chip while Figure 2 shows the flowchart of biometric finger print design.

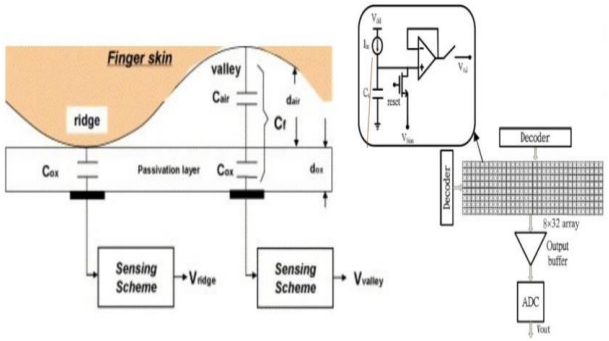


Fig. 1. The theory and architecture behind a capacitive fingerprint scanning chip

2.2 Enrolment / Verification of the Biometric System

This section shows the sequence of Biometric Encryption algorithm as illustrated in Figures 3 – 7.

3 Test, Results and Discussion

3.1 Testing and Evaluation

The model is evaluated using the fingerprint verification/authentication to start examination and to submit examination during the examination period and a series of experiments were performed focusing on the effectiveness, speed and its usability. The usability testing technique is a technique for ensuring that the intended users of the system can carry out the intended task efficiently, with a limited time.

3.1.1 Test for Enrolment Time

The program was tested using 10 different persons; 5 boys and 5 girls, the time it takes for one person fall between 17sec to 35sec, average of 25.4sec per person, as shown in Table 2.

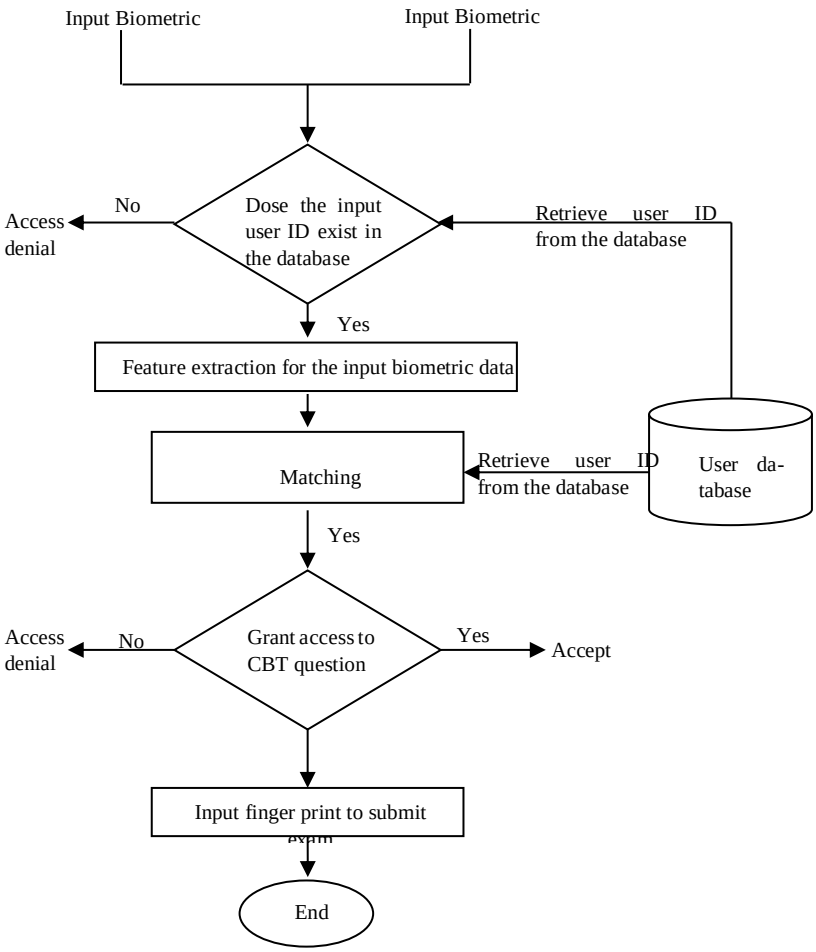


Fig. 2. Flowchart of biometric finger print design

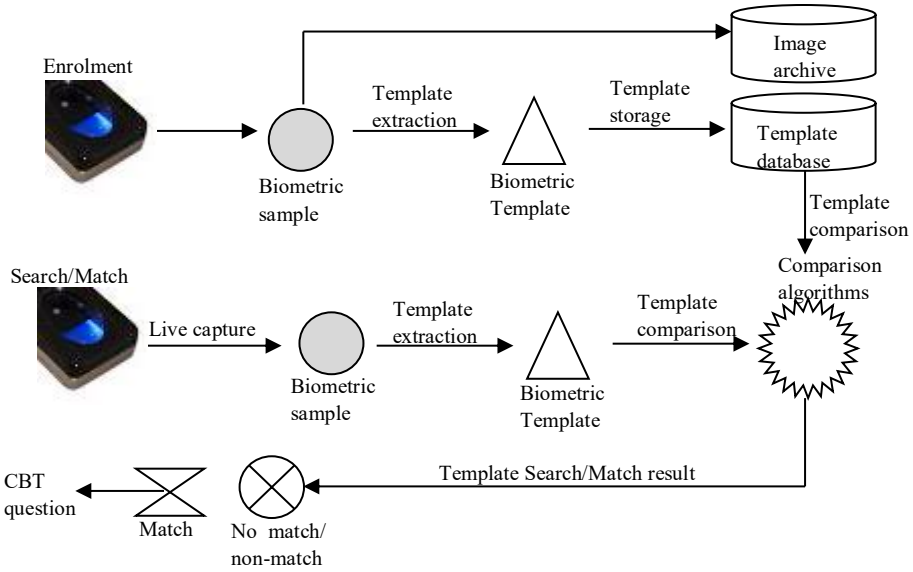


Fig. 3. Overview of the enrolment/verification process for Biometric fingerprint connecting to CBT

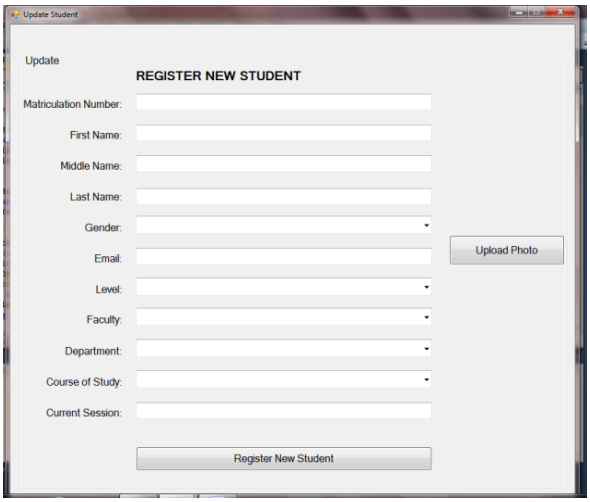


Fig. 4. Biometric fingerprint enrolment for new user sample

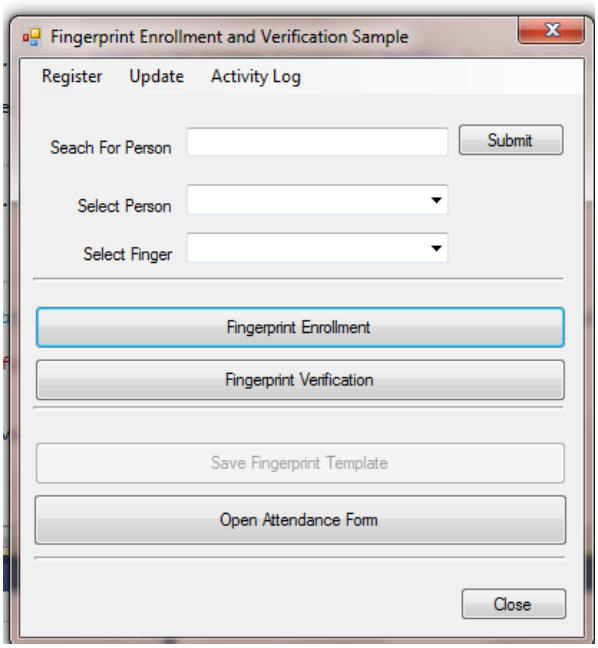


Fig. 5. Biometric fingerprint enrolment/ verification sample

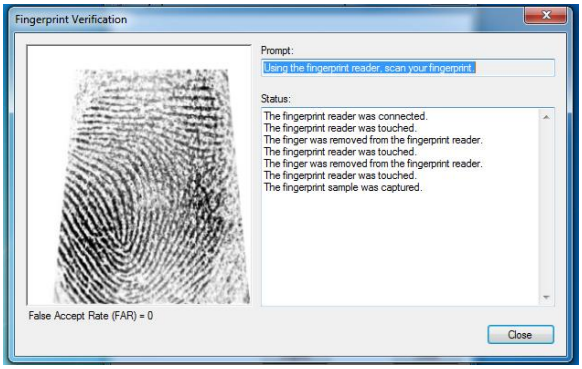


Fig. 6. Biometric fingerprint verification sample

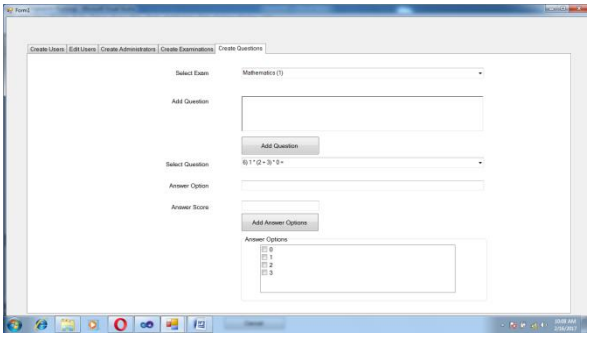


Fig. 7. Computer based test sample

3.1.2 Speed of Verification and Authentication

This test was used to measure the average time it takes to verify/authenticate students' manual with biometric and attendance signature with biometric signature, Table 3 shows the result of manual (average 6.6sec)/biometric verification and authentication (average 1sec) and manual signature (average 11.4sec), one will note that for biometric signature will be generated from the program.

Table 2. Time taken for student to enrol in biometric

Sample	Time Taken (sec)
Joy Osaretin	17
Blessing Agho	20
Abohwoyere Omokhua	33
Osatohawmen Osamuyi	34
Amarachukwu Anajemba	35
Vivian Anokwuru	26
Denoraah Zainab	20
Ashonofor Wilson	24
Chigozie Ebito	21
Eguavoen Iziegbe	24
Total	254
Total Average	25.4

Table 3. Manual/biometric verification and authentication

Sample	Manual verification/ authentication (sec)	Manual nature (sec)	Biometric Sig-verification/ authentication (sec)
Joy Osaretin	7	6	1
Blessing Agho	5	15	1
Abohwoyere Omokhua	5	7	1
Osatohawmen Osamuyi	9	10	1
Amarachukwu Anajemba	8	17	1
Vivian Anokwuru	4	13	1
Denoraah Zainab	9	11	1
Ashonofor Wilson	6	8	1
Chigozie Ebito	10	13	1
Eguavoen Iziegbe	3	14	1
Total Time	66	114	10
Average Time	6.6	11.4c	1

3.1.3 Test of Two-way Authentication with Biometric Authentication

Table 4 shows the time it takes to type in the username and the password in the two-way authentication (average 15.2sec) and the Biometric authentication (average 1sec).

Table 4. Result of Two-way Authentication with Biometric Authentication

Sample	Two-way authentication (sec)	Biometric authentication (sec)
Joy Osaretin	11	1
Blessing Agho	12	1
Abohwoyere Omokhua	18	1
Osatohawmen Osamuyi	18	1
Amarachukwu Anajemba	19	1
Vivian Anokwuru	14	1
Denoraah Zainab	14	1
Ashonofor Wilson	16	1
Chigozie Ebito	13	1
Eguavoen Iziegbe	17	1
Total Time	152	10
Average Time	15.2	1

3.2 Test of program to general requirements

The software developed was also tested and evaluated based on the following criteria;

(1) Ease of Enrolment.

(2) Fast to verify/authenticate.

(3) Adherence to Verification Rules for Examination eligibility.

(4) Fast communication between systems to server

The software was found to satisfy the above criteria given the following observations:

(1) The developed system also ensures that only students who are to sit for exams are allowed access into the examination platform.

(2) It was observed using star connection for the network (LAN) is the best, when one system goes down other will be working.

(3) Student and course attendance report are available to the system administrator, it can be concluded that the developed system effectively addresses the needs of CBC removing malpractice from academic environment with regards to exams and student attendance. The system respond time was 1ms during system to server communication and less then 2ms during authentication; efficiency, and reliability at real time was very good.

4 Conclusion

Security is ongoing process where due care and diligence to protect examination need to be put in place, without security in network systems, most system will be highly exposed to a lot of dangers and threats. Different information technologists have developed several tools, design phases and other techniques to help in the development of standard computer based centre. Most of the technologist had not looked into biometric for authentication and MAC address System. One can recall that why one needs security in the examination to eliminate organised malpractice, a summary of comparison of biometrics is shown in Table 2 Managing the speed of manual verification and authentication with the biometric verification and authentication. The lapses recorded in traditional methods of recording and managing attendance has also be solve by generating attendance from the develop system, to ensure the integrity of such records; biometrics is a tool that cannot be neglected. Fingerprint authentication has thus been tested and proven as a veritable tool in achieving the much needed automation. The result shows that the average time taken per student using biometric and manual verification/authentication register are 1 and 6.6 seconds respectively.

The major strength of the developed system lies in its high scalability and flexibility, by careful examination, it can be inferred that biometric authentication could not only speed up the process of taking exams, attendance but reduce the error rate and produce faster verification/authentication process of student verification/authentication policy required for writing examination in a campus environment. This paper presented a simplified, low cost fingerprint based system solution to the management eliminating malpractice/irregularity in examination. However, it might also be necessary to investigate student through hybridized biometric features like face and iris for better performance.

5 Recommendation

It is highly recommended that the management of Universities take examination/irregularities issues seriously. Biometric verification/authentication promises to provide the global solution with a sound identity management system, which could eliminate malpractice from the University. If biometric security measure is not strictly adhered to in examination in the University and other places for signing in/out, it could pave way for malicious intruders to have access to the University. If examinations are not properly secured and falls into the wrong hands, it could spell doom for such a University. In addition, if security measures are not strictly adhered to, these could pave way for hackers and malicious intruders to have access to examinations.

Conflict of Interest: none declared.

References

Adam Mohammed Saliu, Mohammed Idris Kolo, Mohammed Kudu Muhammad, Lukman Abiodun Nafiu, (2013): 1(1) Internet authentication and billing (hotspot) system using MikroTik router operating system, *Int'l J. of Wireless Communications and Mobile Computing*: pp 51-57.

Adebayo O. and Abdulhamid, S. M. (2014): E- Exams System for Nigerian Universities with Emphasis on Security and Result Integrity. *Int'l J. of the Computer, the Internet and Management*, pp 12,18,1-47.

Beaver Kevin. "Hacking for Dummies®," 4th Edition 2013 by John Wiley & Sons, Inc., Hoboken, New Jersey. Pp 4, 9-13, 27-37, 81-154.

http://www.androidauthority.com/how-fingerprint-scanners-work-670934/

http://www.wisdom24x7.com/

Itakura, Y., Tsujii, S. (2005). Proposal on a multifactor biometric authentication method based on cryptosystem keys containing biometric signatures. *Int'l J. of Information Security*. Heidelberg (4)4, 288

Layton Timothy P. (2007): Information security design, implementation, measurement, and compliance Boca Raton, FL: Auerbach publications. pp 8-13, 40, 62, 68,92, 112.

Luecht, R. M. (2005a). Operational issues in computer-based testing. In D. Bartrum and R. Hambleton (Eds.), Computer-based testing and the Internet. New York: Wiley & Sons Publishing.

Luecht, R. M. (2005b). Some useful cost-benefit criteria for evaluating computer-based test delivery models and systems. *Association of Test Publishers Journal*. Retrieved from www.testpublishers.org/journal.htm Foster, D. (April, 2011). Personal communication.

Nafiu, L. A. (2014): Unpublished Lecture Notes on Net- Centric Computing. Federal University of Technology, Minna, Nigeria.

Thomas Robertazzi. (2012): "Ethernet", Basics of Computer Networking, pp 1-129.

Todorov**,** (2007): "Authenticating Access to Services and Applications", Mechanics of User Identification and Authentication Fundamentals of Identity Management, pp 760.

American Journal of Advanced Research, 2017, 1–1
December, 2017, pages 19-25
doi: 10.5281/zenodo.1227461
http://www.ibii-us.org/Journals/AJAR/
ISBN 2572-8849 (Online), 2572-8830 (Print)

Hierarchical Kernel and Sub-kernels

Francisco Casanova-del-Angel

Instituto Politécnico Nacional. Ciudad de México. México

Miguel Othón de Mendizábal S/N. Col. La Escalera. CP. 07320 Ciudad de México. México

Tel. (5255) 57296000. Ext: 53153

Email: fcasanova49@prodigy.net.mx

Received on August 7, 2017; revised on October 6, 2017; published on October 15, 2017

Abstract

This paper shows the theoretical development of hierarchy by kernels and an algorithm used to obtain an interesting class or partition from a hierarchy. Also shown is the theorem about the *Kernels Optimal Criterion* and how it is expressed as a function of the masses of the points of the vector space and product scale points, the inertia of the cloud formed by those two points or hierarchical nodes, which are called subcores or sub-kernels. The application is made on the terminal efficiency of postgraduate degrees at ESIA, IPN Mexico, along its first 48 years of academic and scientific life and the development of students´ graduation.

Keywords: Hierarchical kernels; sub-kernels; cores; inertia; classes

1 Introduction

From the theoretical standpoint, the purpose of this paper is to analyze, under the precepts of data analysis, the relationship which exists or lies between a hierarchical kernel and the sub-kernels. I also considered, for the construction of the hierarchical classification, the influence that have kernel and sub-kernel concepts at the time of the interpretation of the hierarchical tree.

But why the existence of the purpose? Well, the answer is simple, since one of the major problems that arise at the time of interpretation and validation of results in hierarchical classification is the confidence given to the separation of classes of values in the hierarchical tree, and above the meaning of the height to which are those same classes that build the hierarchy. The height is called level index.

In relation to the application of the theory and interpretation of results on real data, let me explain then what seems to be the classic tedious and gruesome development of education in underdeveloped societies academically and where the accountability is very little practiced.

On December 14, 1961, the General Director of the National Polytechnic Institute Mexico, IPN, presented before to the Technical Advisory Board of the Institute the advantages and rationale to establish the Graduate Section of the Higher School of Engineering and Architecture, ESIA, IPN. On July 16, 1962, registration into postgraduate degrees at ESIA IPN, were formally started, with two master degrees in sciences, with specialty in structures and hydraulics [1]. Courses formally started on Wednesday, August 1, 1962. The syllabi of both master degrees were organized in half-year periods. On December 14, 1966, the School Consulter Technical Committee of the IPN discussed, in the fourth point of its agenda, the creation of six master degrees and a doctorate degree in sciences, among which there was a new master degree in sciences in planning. In 1978, an unprecedented expansion was intended, based on a proposal to create a master degree in sciences for every civil engineering specialty. Despite the boldness of such proposal, four new master degrees in sciences were established: architecture, environmental engineering, geology and soil mechanics. In 1981, mining and oil specialties for the master degree in geology, and architecture and architecture specialty with options for architectonic design and works construction and control were created. In 1983, the structural analysis, steel structures, cement structures, architecture and ports development specialties, for the master degree in hydraulics [1] were created. On June, 1998, through the execution of a General Academic Collaboration Agreement with the Polytechnic University of Madrid, UPM, a joint civil engineering doctorate degree with specialty in environmental hydraulics was established. Such doctorate degree was taught at SEPI ESIA, ALM Unit, IPN, with the support of the Civil Engineering -hydraulics and energy-Department of the UPM. Years later, educational and research institutions must report to the society on the resources provided for the creation and support of these graduate programs. Its usefulness and provided benefits who gave such economic support, since it is never done.

2 Properties of the kernel

A vector space is a set V provided with two operations: the addition of elements of V and the multiplication of elements of V with a scalar.

A mapping T of a vector space V into a vector space W is called a linear transformation of V into W if, for any vectors α, β ∈ V and an arbitrary real number r, the following hold: i) T(α+β) = Tα + Tβ and ii) T(rα) = rT(α). Si $W = V$, the transformation T is often called an operator on V. Associated with any linear transformation T: $V \rightarrow W$, are two very important space in data analysis: the rank space or range denoted by R_T; and the kernel or the null space of T, denoted by N_T, defined in (1) as:

$$N_T = \{T\alpha = 0 \mid \alpha \in V\} \quad (1)$$

We also know that if V is the space on which T operated, we define K_i to be the kernel of the operator T - λ_i; that is, K_i is the subspace of vectors $\alpha \in V$, such that (T + λ_i)α = 0, and so its nonzero members are the Eigenvalues of T that belong to λ_i. [2] pp. 94-121.

3 Theoretical development of hierarchy by kernel

Based on the fact that factorial correspondence analysis represents, on the same graphic, both sets comprising a tabular correspondence arrangement; sets I of individuals and Q of classes defined for each variable J, and that when such must be taxonomies, a rigid class system must be fixed, then the global and spatial vision provided by factorial analysis allows us to establish through some kind of aggregation method, a type of hierarchy of the data under analysis.

The method herein shown is tributary to three options: i) calculation of the distance between elements where factorial coordinates are known; ii) juxtaposition of mass or weight to each element; and iii) calculation of a distance between element classes, depending on an aggregation criterion based on cores. Since our data include factorial values related to Q classes, we shall retain a small number of A cardinality factors, not higher than 75% of factorial data.

Let us define factorial set of values through set: $\{F_\alpha(q) \mid q \in Q \ and \ \alpha \in A\}$, with which it is possible to calculate many tabular arrangements for distances between elements. In our case, we shall introduce the following distance. Let q and q' be two classes of a variable $j \in J$ such that q and $q' \in Q$. Classes q and q' belong to a normed factorial space with a fixed set of coordinates. If $d: F \longrightarrow \mathbb{R}$ then (F, d) is a metric space. Factorial distance between $F(q)$ and $F(q')$ is the addition of lengths of projections of line segment between factorial values on the axes system. This is mathematically expressed as follows [3] and [4]:

$$d^2(q,q') = \| q,q' \|^2 = \sum_{\alpha \in A}(F_\alpha(q) - F_\alpha(q'))^2 \quad (2)$$

Where q and q' are classes of variable $j \in J$, d is the distance between classes, α is the axis, A is the set of axes and $F_\alpha(q)$ and $F_\alpha(q')$ are factorial values of classes. In accordance with the second option of the aggregation method defined, the distance between classes is juxtaposed by inertia λ of the set of dots along axis α, which is represented by the own value related to the corresponding axis, because of this equation (2) may be re-expressed as follows:

$$d^2(q,q') = \| q,q' \|^2 = \sum_{\alpha \in A} \lambda_\alpha^{-1}(F_\alpha(q) - F_\alpha(q'))^2 \quad (3)$$

Where q and q' are the classes of variable $j \in J$, d is the distance between classes, α is the axis, λ_α^{-1} is the inverse of distance between classes on axis α and $F_\alpha(q)$ represents factorial value of class q on axis α [4] and [5]. Once the distance between values has been defined, the diameter index of nodes of classification ν of such hierarchy must be calculated, through:

$$\nu(n) = \frac{f_a * f_b}{f_a + f_b} \| F_\alpha(a) - F_\alpha(b) \|^2 \quad \forall \ n \in Nodo \quad (4)$$

Where a and b are barycenter's of elements of the index, f_a and f_b are the mass in a and b barycenter's, and $F_\alpha(a)$ and $F_\alpha(b)$ are factorial values of a and b barycenter's. In addition, $a \cup b = n$ and $a \cap b = \Phi$.

Every time, the distance between elements that are hierarchized must be recalculated with those to be hierarchized, because of this the following diameter index $\nu(n)$ is:

$$\nu(n) = \frac{f_a * f_b}{f_a + f_b} \| \lambda_\alpha^{-1} F_\alpha(a) - \lambda_\alpha^{-1} F_\alpha(b) \|^2 \quad \forall \ n \in Nodo \quad (5)$$

Where $\nu(n)$ is diameter index, f_a and f_b are masses of a and b barycenter's, $F_\alpha(a)$ and $F_\alpha(b)$ are factorial values of a and b barycenter's, and λ_α^{-1} is the square root of total distance of the A set of dots, along axis α.

Now, from equation (4) it may be seen that the addition of values of diameter indexes is equal to the addition of total distance λ of the set of dots along α axis, that is:

$$\sum_{n \in Nodo} \nu(n) = \sum_{\alpha \in A} \lambda_\alpha \quad (6)$$

Where $\nu(n)$ diameter is indexed and λ_α is the total distance of the set of axes. From equation (5) it may be seen that the addition of the values of diameter indexes is equal to A's cardinality.

$$\sum_{n \in Nodo} \nu(n) = Card\ (A) \quad (7)$$

3.1 The algorithm

Classification algorithm looks for two minimum values of the table of factors of classes of the sub-kernels to be hierarchies.

$$\delta(q,q') = \frac{f_q * f_{q'}}{f_q + f_{q'}} \| F_\alpha(q) - F_\alpha(q') \|^2 \quad \forall \ q,q' \in Q \quad (8)$$

From this aggregation, defined as $k = q \cup q'$, a new partition or kernel of the set of Q classes must be updated making: $\mathcal{P} = Q \cup \{k\} - \{q,q'\}$. Distances between this new element k and q'' are recalculated, showing the following minimum value of the factors table, through formula (4), thus making $\nu(n) = \delta(a,b)$.The minimum of the new table is investigated, aggregated and a new partition is updated below. The above is carried out until there are no more than the two last cores to be added, taking into account that the link is the base set [5] and [6].

Theorem Kernels Optimal Criterion. If aggregation kernels are groups of factors with same cardinality and Ω the space of kernels or cores, the optimal election criterion is:

$$d(L,P) = \sum_{i=1}^{k} d(A_i - P_i)$$

Where L is the total set of kernels or cores, A_i is the ith core containing a certain number of objects of P population.

Demonstration. Let $L = \{A_1, \dots, A_h\}$, $A_i \subset \mathcal{L}$ be the ith kernels or core containing q elements of population. $P = \{P_1, \dots, P_h\}$ is partition of space Ω into k-classes. Let $\mathcal{L}_k$ be the set of kth cores and $\mathcal{P}_k$ the set of partitions of Ω kernels space into classes. $d(A_i, \mathcal{P}_i)$ measures dissimilarities between kernel or core A_i and class $\mathcal{P}_i$. Based on the above, the principal problem is to look for a $L^* \subset \mathcal{L}_k$ and a population $P \subset \mathcal{P}_k$ that minimize d dissimilarity.

Let $d(q_1, q_2)$ be a measure for dissimilarities between couples of individuals or classes. Let us suppose that:

$$d(q_1, q_2) = \sum_{q_1 \in X} \sum_{q_2 \in Y} d(q_1 - q_2)$$

Where X and Y are parts of the set of Ω individuals, then:

$$d(q_2, \{q_1\}) = d(Y, q_1) \quad \text{and} \quad d(\{q_1\}, Y) = d(q_1, Y)$$

In case that kernels or cores are groups of individuals, the algorithm shall be specified, since such is based on choosing two functions: assignation function and representation function. For the assignation function, given the kernels or cores $\{A_1, \dots, A_h\}$, partition $P = \{P_1, \dots, P_h\}$ deducted is defined by:

$$P_i = \{q_1 \in \Omega \mid d(A_i, q_1) \leq d(A_j, q_1) \; \forall \; i, j\}$$

In case of equality, q_1 shall be assigned to the lowest index class. Partitions P thus deducted from L are shown by $P = f(L)$, where f is an application of $\mathcal{L}_k$ in $\mathcal{P}_k$; that is: $f: \mathcal{L}_k \longrightarrow \mathcal{P}_k$, and it is called assignation function.

For the representation function, given partition P, $L = \{A_1, \dots, A_h\}$ kernels or cores are deducted as:

$$A_i = \{q_1 \in \mathcal{L} \mid q_1 \in \{q\} \; wich \; produce \; lowest \; possible \; dissimilarity \; d(q_1, \mathcal{P}_i)\} \quad (9)$$

In order to ensure the unit of A_i, the set of q elements of Ω space minimizing $\sum_{q_1 \in A_i} d(q_1, \mathcal{P}_i) \; \forall \; \mathcal{P}_i \subset \Omega$, exists and is unique. Therefore, the representation function exists. ***QED***

3.2 Sub-kernels

Let a vector space V of W, if it exists $U \subset V$ not empty then U is a vector subspace of V if it complies with the properties given in § 2. Therefore, sub-kernel means a subset $N \subset N_T$. Now that we have seen the principal theorem of hierarchical cores and the implementation of his algorithm, let's see how it is expressed, depending on the masses of the points of the vector space and the scalar product of these points, the inertia of the cloud formed by those two points or hierarchical nodes, which in our case are called sub-cores or sub-kernels forming the principal node of the hierarchy.

Usually the inertia In(g) (or In(h)) part of the cloud N(I) is given by:

$$In(g) = \sum \left\{ r_{ii'} \left\| i_V - i'_V \right\|^2 / \, i, i' \in V \right\} = \sum \left\{ \frac{m_i m_{i'}}{2m_g} \left\| i_V - i'_V \right\|^2 / \, i, i' \in g \right\} \quad (10)$$

In the first part of (10), the double sum includes (Card g)^2 terms (or (Card h)^2 terms). For the proof of (10), it is enough with to replace $i_V - i'_V$ by $(i_V - g_V) - (i'_V - g_V)$ and to develop the square with what you get, when the sums:

$$½ \, \text{In}(g) + ½ \, \text{In}(g) + 0 = \text{In}(g)$$

For the people not familiar with the data analysis, it is understood by class or tax on to the taxonomic division of finite size [7] pp. 94.

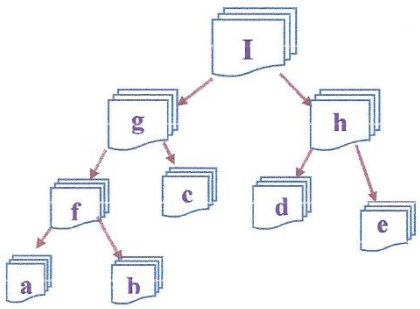


Figure 1. Taxonomic system of sub-kernels.

Finally, to express the inertia of the subspace I fitted with a system of classes or sub-kernels; as shown in Figure 1. The set (a, b, c, d, e, f, g, h) are parts of I that have properties such as a and b, as well as d and e or g and h; among many others, are two to two empty intersections and their union is I, i. e, the set $\{a, b, f, c, d, e\}$ is a partition of I in a number of sub-kernels; which, in the case of Figure 1, are five sub-kernels or classes. In addition is that, for example: $f = a \cup b$, $g = f \cup c$ y $I = g \cup h$.

Given any two parts of I, denoted by $d(a, b)$, the inertia of the point cloud consisting of points a_υ and b_υ with their respective masses m_a and m_b, is possible to write that the indices of diameter are: $\upsilon(I) = d(g, h)$, ….., $\upsilon(f) = d(a, b)$.

In addition, $\upsilon(a) < \upsilon(f) < \upsilon(g) < \upsilon(I)$. With the above, it is possible to express the inertia In(I) depending on the index diameter of the kernel I index; $\upsilon(I)$ and their sub-kernels in the following manner:

$$\text{In}(I) = d(g, h) + \text{In}(g) + \text{In}(h) = \text{N}(I) + \text{In}(g) + \text{In}(h)$$

That mind a classic decomposition of total inertia In(I) in inertias inter kernels of the cloud N(I) and the inertia produced by the addition of the inter kernels of g and h; In(g) + In(h). This last also can be expressed as the inertia associated with centers of gravity of the kernels.

4 Hierarchical Kernel in pseudocode

This algorithm in pseudocode (Figure 4) synoptically describes the operating principle for the production of kernels and hierarchical sub-kernels. The one which, based on the theory developed here, can be implemented in any programming language, or you can make use of commercial software of mathematical statistics.

Table 1. Historical terminal efficiency of master's degrees in sciences.

Master Degree in Science	Number of graduated students	Period	Year of defense of the first specialty thesis	Terminal efficiency annual index
Structures	58	1962-2010	1970	1.20
Hydraulics	51	1962-2010	1975	1.06
Planning	66	1966-2010	1977	1.37
Soil Mechanics	32	1981-2010	1987	0.66
Environmental Engineering	114	1977-2010	1979	2.37
Doctorate degree				
Environmental Hydraulics	5	1998-2010	2000	0.50
Master degree in Engineering				
Structures	3	2009-2010	2010	1.5
Hydraulics	5	2009-2010	2010	2.5
Planning	5	2009-2010	2010	2.5
Geotechnics	0	2009-2010	-	0
Environmental Engineering	10	2009-2010	2010	5

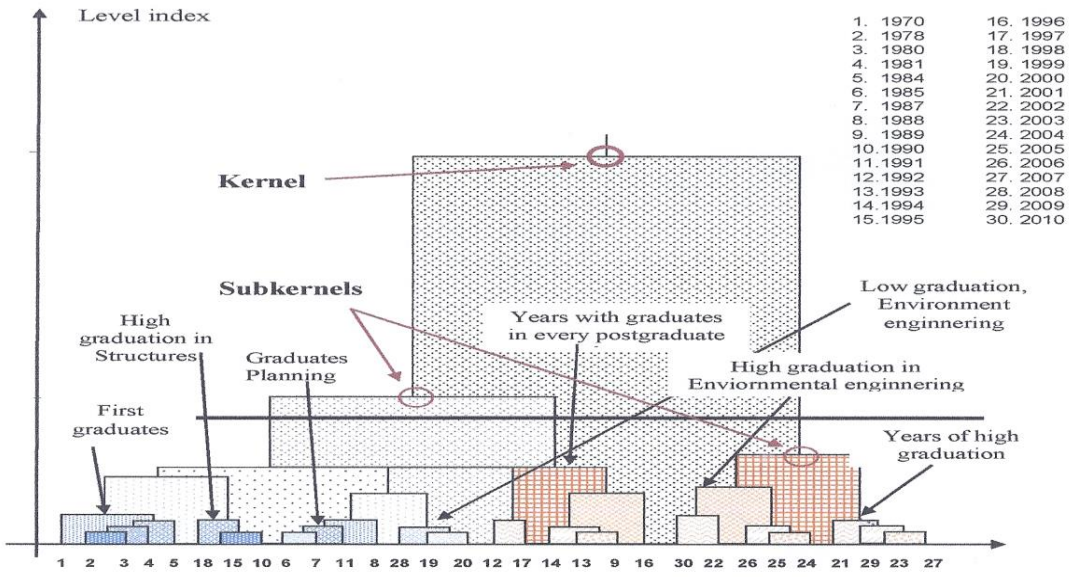


Figure 3. Hierarchical classification of terminal efficiency of postgraduate degree ESIA IPN, Mexico with hierarchical kernels theory.

Table 2. Correlations matrix

	1	2	3	4	5	6	7
1	1.000						
2	0.652	1.000					
3	0.377	0.554	1.000				
4	0.564	0.656	0.695	1.000			
5	0.630	0.502	0.431	0.535	1.000		
6	-0.229	-0.018	-0.044	-0.035	-0.60	1.000	
7	-0.159	-0.098	-0.202	-0.178	-0.086	0.526	1.000

Where 1: Structures, 2: Hydraulics, 3: Planning,
4: Environmental engineering, 5: Soil mechanics,
6: Hydrocarbons administration and 7: Geotechnics.

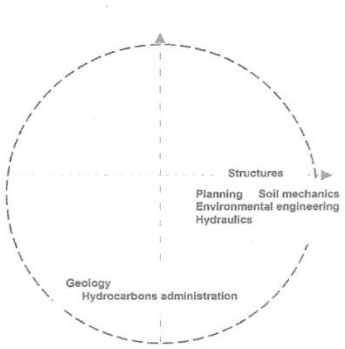


Figure 2. Correlation Circle of first principal component *vs* second principal component.

//For the preparation of the table classes//
For $i \leftarrow 0$ **up to** n **Make**
  Input $(r_1 \leq j_1 \leq r_2)$ ***up to*** $(r_{m-1} \leq j_m \leq r_m)$ **Make**
  For $j_q \leftarrow (r_1 \leq j_1 \leq r_2)$ ***up to*** $(r_{m-1} \leq j_m \leq r_m)$
    Instruction $k_{IQ} = k(i, j_q) \quad \forall\ i \in I, j \in J, q \in Q - class$
    Write $k(i, j_q)$
  End for r_m
End for i
//The calculation of correlations or *degree of association* among variables in class, was carried out from usual Euclidian distance//
For $i \leftarrow 1$ **up to** n **Make**
  Input j, j' **and define distance**
    Instruction $d^2(j, j') = \sum_{i=1}^{n} x_{ij}^2 + \sum_{i=1}^{n} x_{ij'}^2 - 2\sum_{i=1}^{n} x_{ij}\, x_{ij'}$ **Make**
    Write and Draw $d^2(j, j')$
End for i
// Calculation of the distance factor data in class//
For $\alpha \leftarrow$ 1 **up to** 7 **Make**
  Instruction $d^2(q, q') = \| q, q' \|^2 = \sum_{\alpha \in A} (F_\alpha(q) - F_\alpha(q'))^2$ **Make**
  Instruction λ_α^{-1} **Make**
  Instruction $d^2(q, q') = \| q, q' \|^2 = \sum_{\alpha \in A} \lambda_\alpha^{-1} (F_\alpha(q) - F_\alpha(q'))^2$ **Make**
  Write $d^2(q, q')$
End for α
//Built of the class hierarchy or aggregation criteria and draw level index//
For $\alpha \leftarrow$ 1 **up to** 7 **Make**
  Input q **and** q'
  Instruction f_q **and** $f_{q'}$ **Make**
  Instruction $\delta(q, q') = \frac{f_q * f_{q'}}{f_q + f_{q'}} \| F_\alpha(q) - F_\alpha(q') \|^2 \quad \forall\ q, q' \in Q$ **Make**
  Write and Draw $\delta(q, q')$
End for α
//Application of theorem of kernels and sub-kernels//
For $i \leftarrow 1$ **up to** k **Make**
  Instruction $d(L, P) = \sum_{i=1}^{k} d(A_i - P_i)$ **Make**
  Write and Draw $d(L, P)$
End for i
// Reading and interpretation of the hierarchical kernels//
End of Hierarchical Kernels and Sub-kernels
End of Procedure

Figure 4. Pseudocode

5 Application

Currently, one of the criteria used to assess the functioning of academic and research activities is terminal efficiency, as one of the principal indicators showing the achievements of the corresponding education institution. Since the School of engineering and architecture. Unit Adolfo Lopez Mateos of the Polytechnic Institute National. Mexico, is has been one of the schools of civil engineering with more students in Mexico, it is very important to know its terminal efficiency, both for licentiate and postgraduate degrees, see [8] and [9]. On the top of the table, the number of graduates for each master's degree in sciences that is officially known up to 2007, throughout 48 years, is shown. On the bottom of table 1, the terminal efficiency of the master's degree in civil engineering up to date, which substituted the five previous ones in 2007 is shown. The institution does not update your data automatically, due complicate administrative processes.

5.1 Correlation of terminal efficiency

The calculation of correlations or *degree of association* among variables was carried out from usual Euclidian distance $d(j, j')$ among variables j and j'; that is: $d^2(j,j') = \sum_{i=1}^{n} x_{ij}^2 + \sum_{i=1}^{n} x_{ij'}^2 - 2\sum_{i=1}^{n} x_{ij}\, x_{ij'}$. Since general terms of normed analysis in general terms in real space of dimension p, $\Re^p$, are points x_{ij} we have that: $\sum_{i=1}^{n} v_{ij}^2 = \sum_{i=1}^{n} x_{ij'}^2 = 1$. Every point-variable is on a sphere with radius 1 and center on the origin of principal axes, which the correlation coefficient $c_{ij'}$ among variables j and j' is: $\sum_{i=1}^{n} x_{ij}\ x_{ij'} = c_{ij''}$. The correlation matrix is shown as Table 2.

Best correlated master degrees in sciences are: environmental engineering planning/hydraulics, and structures/ hydraulics, Figure. 2. It must be remembered that, if two meteorological variables are *strongly correlated,* they are near from each other ($c_{ij'} = 1$) or, on the contrary, as far from each other as possible ($c_{ij'} = -1$), in accordance with linear relationship linking them is direct or inverse, and that when $c_{ij'} = 0$ they are considered at an average distance or that variables j and j' are orthogonal.

5.2 Factorial Correspondence Analysis of Gross Data

The factorial method chosen to describe data under study is the Factorial Correspondence Analysis, FCA, since it allows the direct search of simultaneous representation of sets under study I years of graduation and J master degrees in the sciences [3]. The FCA applied on gross data K_{IJ} has the following factorial characteristics: variances on the first five principal axes or own values are: $X_1 = 3.3186$, $X_2 = 1.4787$, $X_3 = 0.7817$, $X_4 = 0.4838$ y $X_5 = 0.4223$, while the inertia percentages explained by such axes are, respectively: 47.4%, 21.1%, 11.2%, 6.9%, and 6.0%. Principal axes are well defined. The first includes master's degrees in sciences in environmental engineering, hydraulics, and structures. The second principal axe includes the master's degrees in science that did not belong to this school of engineering for a long time, hydrocarbons administration, economy and geology while the third axe is planning.

Figure 3, shows the hierarchy of relationship between the years of graduation of master degrees in sciences. In the upper-right corner of Figure 3 are the years or periods of analysis of available information. They are the years that contain record of students graduating in these graduate programs, which at the same time are the classes that define the hierarchy. Reading and interpretation is based on the value of hierarchical level index, shown on the left of the dendrogram, such being understood as the consecutive order of values from the product of the weight of the class under analysis and its diameter (distance $d(i, i')$ is the diameter of the smallest part of a hierarchy containing both i and a i') [1].

6 Conclusions

This work is presented in accordance with its development. The theory developed on hierarchical cores is shown, where the method shown is tributary to three options: i) calculation of distance between elements where factorial coordinates are known; ii) juxtaposition of mass or weight to each element; and iii) calculation of a distance between element classes, depending on an aggregation criterion based on hierarchical cores. From the point of view of the theory developed, it may be seen that from various starting points, the problem of looking for stable classes may be resolved. Starting points may be chosen by the user, with the help of a hierarchical classification. The theorem demonstrated and called *Cores Optimal Criterion Theorem* allows to implement f *and* f^{-1} functions from a kth core randomly estimated with the algorithm. In relation to the application of the theory, it is possible to say that the hierarchical dendrogram built is formed by three branches, whose interpretation is absolutely congruent with knowledge on the topic.

To achieve the optimal terminal efficiency of the Section of Postgraduate Degrees, a real connection between professor and student must be fostered, in order that information moves in both ways, since a lot of students, along their lives, carry out professional practice highly contributing to the technological and scientific progress, which, together with professors as knowledge guides, may yield significant progress. As a result of the analysis carried out in this work, it must be noticed that one of the areas of knowledge of the postgraduate degree students are more interested in are environmental engineering and planning, offering the highest number of graduates in such specialties.

Acknowledgements

This document was developed with part of the time devoted to the IPN-SIP 20171058 research project.

Conflict of Interest: none declared.

References

[1] Catálogo de los estudios de posgrado ofrecidos en el IPN. (1984). Dirección de Graduados e Investigación. Instituto Politécnico Nacional. México.

[2] Benzécri, J P and Benzécri F. (1980). *Pratique de L´Analyse des Données. 1. Analyse des Correspondances. Exposé Élémentaire*. Dunod Bordas, Paris. ISBN: 2-04-011227-8

[3] Casanova del Angel, F. (2001). *Análisis multidimensional de datos*. Editorial Logiciels, ISBN: 970-92662-1-7. Mexico.

[4] Marion M and Signonello S. (2011). From Histogram Data to Model Data Analysis. B. Fichet *et al* (eds.). Classification and Multivariate Analysis for Complex Data Structures,

Studies in Classification Data Analysis, and Knowledge Organization. Doi: 10.1007/978-3-642-13312-1_38. Springer-Verlag Berlin Heidelberg.

[5] Diday E and Noirhomme M. (2008). *Symbolic Data Analysis and the SODAS software*. Wiley. ISBN 978-0-470-01883-5, 457 pages.

[6] Diday E. (2008). *Spatial classification*. DAM (Discrete Applied Mathematics) Volume 156.

[7] Benzécri J. P, Benzécri F, Bellier L, Bénier B, Blaise S, Bourgarit Ch, Brian, J. P, Cazes P, Dreux Ph, Escofier B, Fénelon J. P, Forcade J, Giudicelli X, Grosmangin A, Guibert B, Hassan A R, Hathout A, Jambu M, Kamal I. H, Kerbaol M, Lacoste A, Lacourt P, Laganier J, Lebesux M O, Lechat J, Le Chappelier M, Lenoir P, Mahé J, Mann C, Marano Ph, Masson M, Moitry J, Müller J, Nakhlé F, Piétri M, M, Richard J F, Rousseau R, Roux G & M, Salem A, Sandor G, Stérpan S, Tabet N, Thauront G, Volle M, Yagolnitzer E y Zloptowicz M. 1976. *L´Analyse des Données 1. La Taxinomie*. Ed. Dunod Paris. ISBN: 2-04-003316-5.

[8] Casanova del Angel, F. 2010. Structure géométrique des distances hiérarchique No. 41, pp. 27-47. Revue MODULAD. ISSN: 1769-7387.

[9] Casanova del Angel, F. 2012. Terminal Efficiency Analysis of a Postgraduate Degree, Under the Theory of Hierarchical Distances Geometric Structure. The 2012 International Conference on Frontiers in Education: Computer Science & Computer Engineering. Las Vegas, Nevada USA. July 16-19, 2012. ISBN: 1-60132-212-7, 1-60132-213-5 (1-60132-214-3), pp. 143-147.

American Journal of Advanced Research
(Quarterly, Started in 2017)
Vol. 1, No. 1, December, 2017
Price: $80

Published and printed by Institution of Business Intelligence Innovation
P. O. Box 218241, Houston, TX 77218, USA
Phone: 1-281-619-5661
E-mail: contact@ibii-us.org
Website: http://www.ibii-us.org

ISSN 2572-8830
9 772572 883008