International Journal of Engineering & Technology
CENTRE OF PROFESSIONAL RESEARCH PUBLICATIONS
ISSN 2049-3444

International Journal of Engineering and Technology Volume 5 No. 11, November, 2015

Investigation of Computer Generated Clusters on Square Lattice Using Box Dimension Characterization

Salau T.A.O,Hashim M.K, Abayomi K
Department of Mechanical Engineering, Faculty of Technology, University of Ibadan, Ibadan, Nigeria

ABSTRACT

The geometry of fractals and the mathematics of fractal dimension have provided valuable tools for variety of scientific applications. This study modelled a square lattice on 2-dimensional Euclidean plane, populated it with Boolean matrix, labelled it with Hoshen-Kopelman (HK) algorithm and determined geometric variation of five largest clusters if any by estimating their Average Estimated Fractal Dimensions (AEFD) at different scales of resolution and occupation probabilities. The randomly generated matrices according to specified occupation probabilities within the square lattice were labelled, counted and the number of cluster existed within the lattice were identified. The average box counting dimension was obtained by implementing the least square regression procedures on the number of boxes counted per different cluster at different scales of observation across the clusters. The (AEFD) of first five largest clusters increases when the occupation probabilities increases. However, the fractal dimensions of the first largest cluster dominated and maintained a steady value beyond the critical probability (0.593) while the fractal dimension of the remaining four largest clusters started decreasing rapidly for all occupation probabilities above the critical probability. All the five clusters enjoy the same complex degree of geometrical characteristics before the reach of critical probability

Keywords: *Percolation, Threshold Probability, Computer Generated Cluster, Box Counting Dimension, Fractal Dimension, Occupation Probability.*

1. INTRODUCTION

The theory of percolation has enjoyed tremendous engineering applications in the past two decades. It has been used to explain and model a wide variety of phenomena that are of industrial and scientific value. Its application examples included characterization of porous materials and reservoir rocks (Zhou, 2003), fracture pattern and earthquakes in rocks, calculation of effective transport properties of porous media (Deckymn, 1994) permeability, conductivity, diffusivity, etc., groundwater flow, polymerization and gelation, biological evolution(Antonio, et al, 2012), galactic formation in the universe, spread of knowledge, forest tree and many others. These wide applications motivated this study with focus on fractal geometric property of cluster, this refer to a term used in percolation theory which describe a collection of neighbours that obeys certain rules (Murphy, 2003).

In addition, (Stauffer and Aharony, 1992) defined percolation as a theory that deals with the number and properties of clusters. According to (Christensen, 2002), a cluster is a group of nearest neighbouring occupied sites, each of the sites in a lattice is randomly occupied with probability p or empty with probability $(1-p)$ independent of the status (empty or occupied) of any other sites in the lattice. Site percolation involves the population of sites randomly with occupational probability which result into clusters formation by the adjoining neighbours within the lattice (Michael, 2008).

(Hackett, et al, 2011) introduced a model for the creation of a class of random networks with nonzero clustering. Within this model the degree of distribution and clustering spectrum of a network were prescribed, and as such could be fitted to given real-world data. There, an analytical approach was presented to site percolation in these networks. Theoretical predictions for the critical site occupation probability and the fractional size of the giant connected component were shown to match well with numerical simulations on both real and synthetic networks. They also demonstrated the application of their approach to Newman's triangle-based model of clustered random networks.

(Wilson, et al, 2013) examined the advantages of using ImageJ, a free open source software for comparative fractal analysis of the 2013 November 5 multiple solar eruption images in the Solar Dynamics Observatory Database. The results suggested observable trends in fractal dimension of the brightness signal before an eruption in the AIA-13.1 nm hot channel which was estimated using box counting method.

(Damiean, 2009) applied percolation theory to financial modelling using the java programming language. The complete implementation of the model was based on the Count-Bouchand model. His finding was that the Count-Bouchand model was capable of generating results distributed as both a Gaussian and a power law in the activity of percolation probability.

(Hamideh, 2001) explored quantitative analysis of partially recrystallized and fully recrystallized microstructures of hot-rolled 7050 Aluminum alloy. He used Matlab program to detect and extract the desired grain boundaries in both fully and partially recrystallized microstructure of aluminum alloy while removing all other feature and applied fractal dimension using box counting dimension. He concluded that at a constant resolution, increasing the number of iterations results in a fractal dimension closer to the mathematical value, while for each Koch curve higher image resolution results in less deviation from the theoretical fractal dimension.

In the field of medicine, (Dharmanna, et al, 2013) employed fractal dimension analysis as diagnostic parameter to detect glaucoma in patients than clinical parameter cup to disk ratio using box counting method and semi-variance method. He found out that fractal dimension could be used as diagnostic parameter instead of cup to disk ratio in the future.

Furthermore, (Reishofer, et al, 2012) investigated the patients with anteriovenous malformation (AVM) with fractal analysis. Box counting method and Minkowski dimension was adopted to quantitatively measure cerebral complexity. It was found out that fractal dimension is a sensitive parameter to capture vascular complexity in patient with AVM.

(Salau and Ajide, 2012) investigated on some selected surfaces that were of engineering interest. The surfaces were sectioned with random by selected sectional plane, while the section image or numerical integration with the use of fractal disk dimension method to characterize the resulting image from simulated surfaces. The estimated disk dimension was obtained by implementing the least square regression procedures on optimum disks counted at corresponding scales of observation. It was concluded that the estimated disk dimension of selected surfaces were close to smooth surfaces.

(Iftekharaddin, et. al, 2002) used fractal analysis to identify and detect tumor in brain magnetic response (MR) images. Magnetic response images typically have a degree of noise and randomness associated with the natural random nature of structure. This research adopted the use of existing fractal based techniques i.e. box counting method and proposed three modified algorithms. Piecewise-threshold box counting (PTBC), Piecewise modified box counting and piecewise-triangular prism surface area were respectively utilized to detect and locate tumor in brain. Cumulative histogram versus fractal dimension were employed for PTBC method, a piecewise fractal dimension computation was also employed for both PTBC and PTPSA methods respectively.

Having gone through some of the literatures in which fractal dimension methods was adopted, the present study required the estimation of fractal dimension of five largest clusters generated at different scales of resolutions and probabilities, box counting dimension is the most appropriate to estimate the fractal dimension because of the grid sizes of the lattice sites in which cluster could be generated.

2. THEORY AND METHODOLOGY

Cluster generation involves site percolation, percolation theory describes how the number and properties of clusters vary with the occupational probabilities and it is independent of others. Various lattice sizes exist in a cluster analysis, square lattice was employed in this research and each site was either occupied or empty. The occupancy and emptiness of each site depend on the occupation probability which premise on the random numbers, while the random numbers generated were real numbers, it was converted into Boolean numbers by setting certain range of values. Any random values less or equal to chosen probability will be assigned one, and other values will be assigned zero.

Therefore, occupied site was given the value of one, and the empty site was given the value of zero. This procedure was carried out using Fortran 90/95 programming language.

Hoshen-Kopelman (HK) algorithm is a simple algorithm for labelling and counting the number of clusters on a grid of square or rectangular network of cells where each cell may be occupied or emptied. The Hoshen-Kopelman is an effective means of identifying clusters of adjoining cells. However, the HK is really just a distinct application of the labelling the clusters within the lattice. The HK has two stages of implementation:

The first stage involved using the HK algorithm is to scan through the grid looking for occupied cell. To each occupied cell it is expected to assign a label corresponding to the cluster to which the cell belongs. If the cell has zero occupied neighbour, then it is assigned to the cluster label it has not been used (it is a new cluster). If the cell has one occupied neighbour, it is assigned to the current cell the same label as the occupied neighbour (they are part of the same cluster). If the cell has more than one occupied neighboring cell, then the lower-numbered cluster label of the occupied neighbour is selected to use as the label for the current cell. However, if these neighbouring cells have differing labels, it must be noted that these different labels correspond to the same cluster. Cluster, the current cell is labelled the value of the occupied cells.

The second stage consists of spanning the lattice a second time, finding and updating the cluster labels (to ensure that each cluster is represented only by one label). Immediately after the first scan, subroutine was coded to link the cluster together. Four subroutines were called to find and update the lattice. The first subroutine was named check-down where the cell was spanned downward, the current cell was compared to its neighbour and the minimum value was chosen between the current cell and its neighbour and this process continued to the end of the column. The second subroutine was tagged check-right where the current cell was compared to its neighbour at its right hand, the minimum value was selected between the current cell and its neighbour, and this process continued to the end of the row. The third subroutine was appended check-top where the column was spanned upward, each current cell was compared to its neighbour and the minimum cluster was chosen and this process continued up the column. The last subroutine was called check-left where the row was spanned through from right to left and current occupied cell was compared to its neighbour and the minimum value was chosen and the process continued towards the left side.

3. CLUSTER IDENTIFICATION AND FRACTAL DIMENSION

When the occupied cells in the lattice has been updating into correct labelling, there was a need to count the number of clusters within the lattice, and each connected occupied sites in the four cardinal directions but not in diagonal directions were counted to known the number of occupied cells that formed a cluster which represent cluster sizes.

One of the methods used in estimating fractal dimension is fractal box dimension in which the fractal properties could be obtained from power related function given by this relationship. (Salau and Ajide, 2012).

$$Y \;\; \alpha \;\; X^{D}$$

3.1

Proportional related equation (3.1) can be re-written as in equation (3.2) below

$Y = KX^D$

3.2

Where Y = Number of occupied site pertaining to cluster in consideration in the lattice in 2-dimensional Euclidean space.
X = Scale of occupancy of the cluster.
D = Fractal dimension. This will be referred in this study as Estimated Box Dimension for Box counting method.
K = Constant of proportionality. Taking the logarithm of equation 3.2

$\log Y = \log K + \log X^D$

3.3

$\log Y = D\log X + \log K$

3.4

Comparing equation (3.4) to the equation of a straight.

$y = mx + c$

3.5

Using linear regression analysis, fractal dimension can be estimated as

$$D = \frac{1}{n\sum x_i^2 - (\sum x_i)^2}\left(n\sum x_i y_i - \sum x_i \sum y_i\right)$$

3.6 D is the estimated fractal box dimension.

4. ESTIMATING FRACTAL BOX DIMENSION OF A CLUSTER

When a square lattice was populated with certain occupational probability, HK algorithm would then be employed to label the clusters and count the number of clusters within the lattice. Afterwards, each cluster could be identified and ready to calculate its fractal box dimension. Lattice is a 2-dimensional Euclidean space containing row and column $L(i,j)$ *where i and j* represent row and column respectively. When a cell was taken as the centre $a(i,j)$, where $a(i,j)$ represent a cell or site in the chosen cluster, the scales were represented by a square matrix formed with respect to the chosen centre. For example, the first scale formed a loop starting with $(i-1, j-1)$ and end with $(i+1, j+1)$ which formed a 3x3 matrix with $a(i,j)$ as a centre, so the first scale is 3.
Consequently, the second scale increased by two to the right and decrease by two to the left which invariably formed a 5x5 matrix with the same centre. Therefore, the scales establishes an arithmetic progression (3, 5, 7, 9…) as the number of row and column increases. The number of occupied sites within the scale would be counted against each scale. The process will continue until the member of the chosen cluster did not appear again in the scale. The scales were represented by X and the number of occupied sites of selected cluster was represented by Y, the natural logarithmic value of Y would then be plotted against the natural logarithmic value of X, the best fit will be plotted using a least square regression analysis, the slope or gradient obtained from the graph indicated fractal box dimension.

5. RESULTS AND DISCUSSION

The fortran code developed was prompted to run in other to illustrate how the lattice was populated, a 10X10 square lattice was considered with random number generator with seed value of 1234 at initial probability of 0.593 which was shown in Table.

Table 1: Square lattice 10X10 at occupation probability of 0.593

1	1	1	0	1	1	1	1	1	1
0	1	1	1	0	0	0	1	1	0
1	1	0	1	1	1	0	1	0	1
0	1	1	1	1	1	1	1	0	1
0	0	0	0	1	1	1	0	1	1
1	0	0	1	0	1	0	1	1	0
0	1	1	0	1	1	1	1	0	1
1	0	1	1	1	0	1	0	1	1
1	0	0	1	1	1	1	1	1	1
1	0	0	1	1	1	1	1	1	0

The lattice was labelled using the HK algorithm to produce good and bad label as shown in Table 2

Table 2: Good and bad label of 10X10 lattice

1	1	1	0	2	2	2	2	2	2
0	1	1	1	0	0	0	2	2	0
3	1	0	1	1	1	0	2	0	4
0	1	1	1	1	1	1	1	0	4
0	0	0	0	1	1	1	0	5	4
6	0	0	7	0	1	0	8	5	0
0	9	9	0	10	1	1	1	0	11
12	0	9	9	9	0	1	0	13	11
12	0	0	9	9	9	1	1	1	1
12	0	0	9	9	9	1	1	1	0

The good and bad label shown in Table 2 was later turned into good label by subroutine call in the Fortran code developed as represented in Table 3

Table 3: Good label of 10X10 square lattice

1	1	1	0	1	1	1	1	1	1
0	1	1	1	0	0	0	1	1	0
1	1	0	1	1	1	0	1	0	1
0	1	1	1	1	1	1	1	0	1
0	0	0	0	1	1	1	0	1	1
6	0	0	7	0	1	0	1	1	0
0	1	1	0	1	1	1	1	0	1
12	0	1	1	1	0	1	0	1	1
12	0	0	1	1	1	1	1	1	1
12	0	0	1	1	1	1	1	1	0

Each cluster that appeared in the lattice has corresponding number of occupied site which represented their sizes. Cluster label and their sizes in the Table 3 was depicted in the Table 4.

Table 4: Cluster labels and their sizes

Cluster label	cluster size
1	63
6	1
7	1
12	3

Number of clusters were denoted by the number of cluster sizes occured in the lattice, if two clusters have same size, it will be regarded as a cluster as it was highlighted in Table 4.

Table 5 showed the cluster sizes and their corresponding number of occurrences.

Table 5: Clusters sizes and their corresponding number of occurrence

cluster sizes	number of occurrences
size 1	2
size 3	1
size 63	1

In addition, at each particular occupation probability, ten iteration were taken as trials and the average of five largest clusters were estimated. Table 6 showed the summary of average fractal dimensions of the five largest clusters at 0.593 occupational probability.

The result of fractal dimension of the largest five clusters lattice depends directly on occupation probabilities and the number of rows and columns specified when the program was being prompted to run and where user is expected to supply those input parameters. Different lattice sizes were run starting from 20X20, 50X50, 100X100, 150X150 and 200X200 with occupational probabilities of 0.001 to 1.000. The results were shown in figure 1, figure 2, figure 3, figure 4 and figure 5 respectively.

For 20X20 lattice size in Figure 1, cluster 2, cluster 3, cluster 4 and cluster 5 decreased rapidly beyond probabilities 0.588, 0.536, 0.510 and 0.50 respectively.

Table 6: Summary of fractal dimension of five largest clusters at probability of 0.593

Trial no	cluster 1	cluster 2	cluster 3	cluster 4	cluster 5
1	1.3	1.073	0	0	0
2	1.249	1.239	0.704	0	0
3	1.431	0.282	0	0	0
4	1.263	0.971	0	0	0
5	1.263	0.603	0	0	0
6	1.357	0.621	0	0	0
7	1.341	1.089	0.704	0	0
8	1.007	0.857	0.83	0.759	0
9	1.296	1.209	0	0	0
10	1.423	0.8	0	0	0
Total	1.293	0.874	0.224	0.076	0

For 50X50 lattice size in Figure 2, cluster 2, cluster 3, cluster 4, cluster 5 decreased rapidly beyond probabilities of 0.62, 0.578, 0.579 and 0.577 respectively.

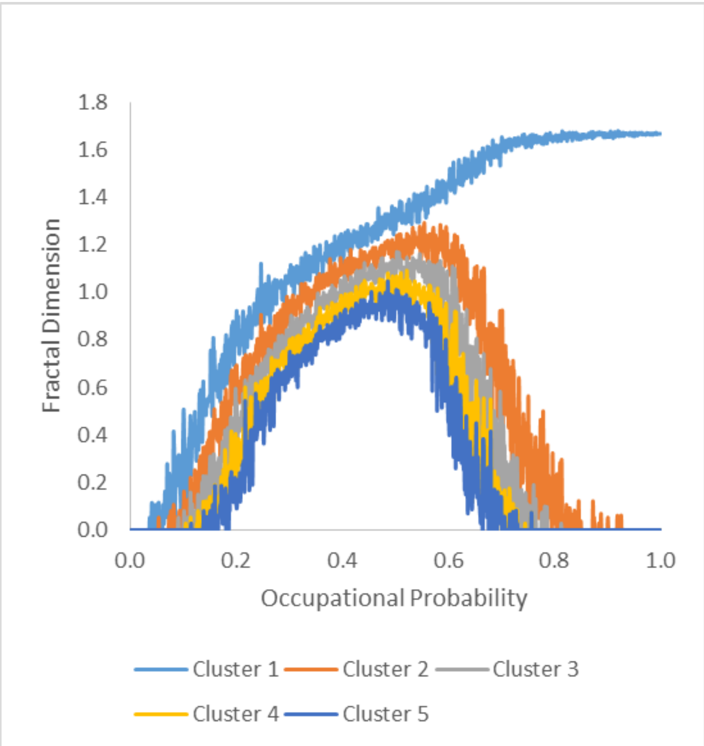

Figure 1 Fractal dimension of 20X20 square Lattice.

For 50X50 lattice size in Figure 2, cluster 2, cluster 3, cluster 4, cluster 5 decreased rapidly beyond probabilities of 0.62, 0.578, 0.579 and 0.577 respectively.

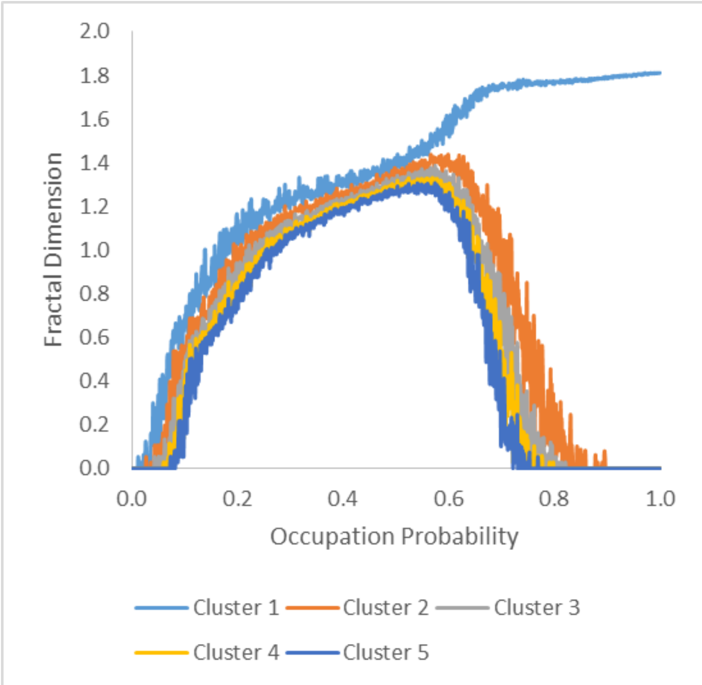

Figure 2 Fractal dimension of 50X50 square lattice

For 100X100 lattice size in Figure 3, cluster 2, cluster 3, cluster 4, cluster 5 decreased rapidly beyond probabilities of 0.667, 0.660, 0.665 and 0.669 respectively.

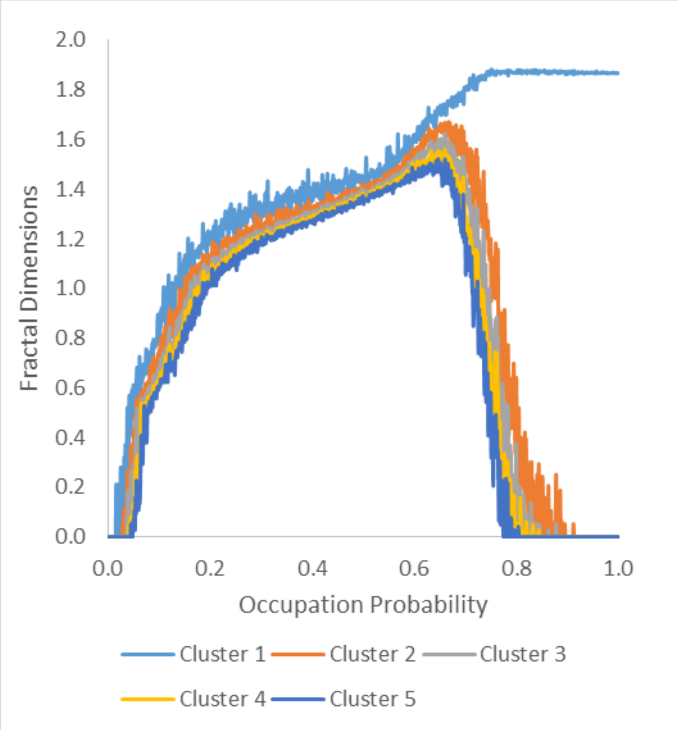

Figure 3 Fractal dimension of 100X100 square lattice.

For 150X150 lattice size in Figure 4, cluster 2, cluster 3, cluster 4, cluster 5 decreased rapidly beyond probabilities of 0.679, 0.668, 0.671 and 0.671 respectively.

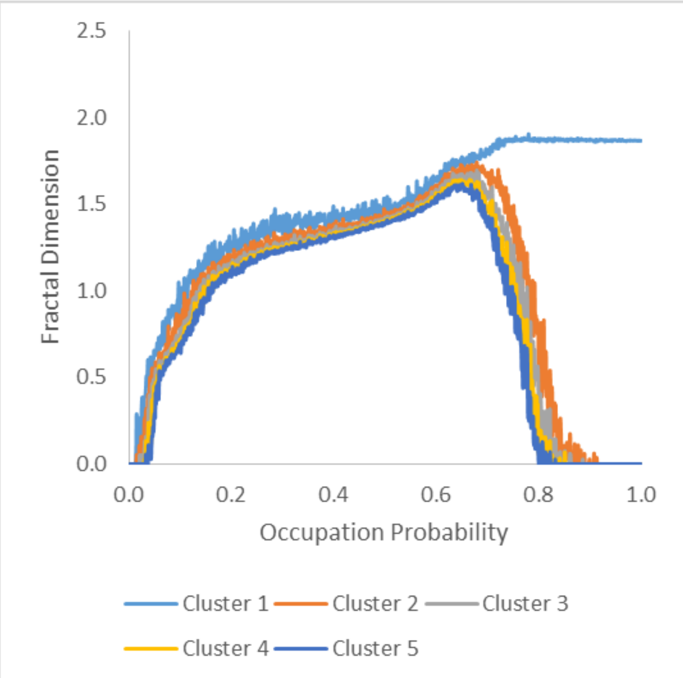

Figure 4 Fractal dimension of 150X150 square lattice

For 200X200 lattice size in Figure 5, cluster 2, cluster 3, cluster 4, cluster 5 decreased rapidly beyond probabilities of 0.690, 0.661, 0.661 and 0.676 respectively.

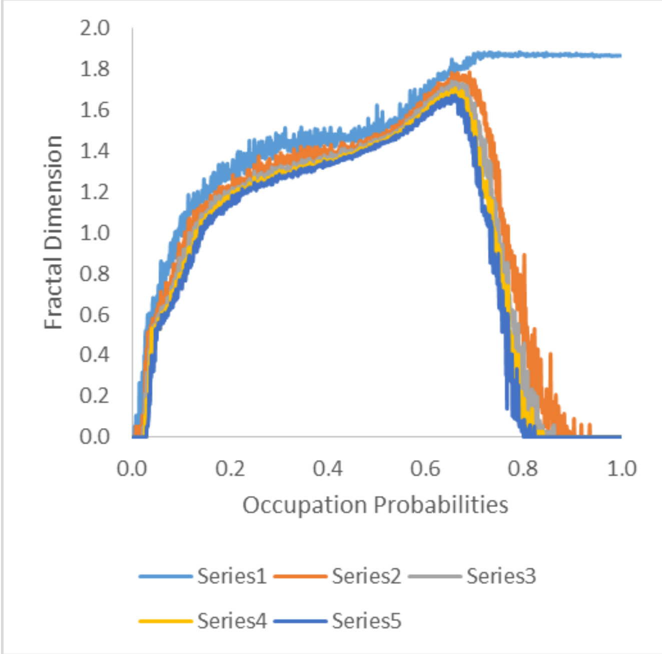

Figure 5 Fractal dimension of 200X200 square lattice

As the lattice size is increasing, the fractal dimension at which the leading cluster i.e. cluster 1 attains a constant value is relatively increased and also the occupation probabilities at which the remaining clusters start decreasing rapidly are also increasingly relatively, but the limitation actually lies in the number of centres taken in estimating the fractal dimensions, when the number of centres taken is low, the execution time will be short but when the number of centre is high, the execution time will be longer and virtually the program take larger percent of computer memory. Basically, there is a limitation to the size of the lattice with respect to the number of centres taken when running the program.

6. CONCLUSIONS

In all the largest five clusters that were considered in this research, the fractal dimensions of the largest five clusters have the same geometric characterization before the reach of critical probability irrespective of size of lattices involved while the box fractal dimension of the last four largest clusters decrease rapidly for all the occupation probabilities above the threshold value of 0.593 according to the literature. As the lattice increases in size, the point of inflexion move away from the percolation threshold.

REFERENCES

Antonio N, Sara U & Vittorio C (2012) Fractal Dimension Estimation Methods for Biomedical Images. Department of Occupational Health and Safety, Medical Physics. Bambino Gesu Children's Hospital. Rome. Italy.

Christensen, K. (2002) Percolation Theory, Imperial College, London Prince Consort Road SW7 2BW United Kingdom.

Damien, D. (2009) On Percolation Theory for Agent-Based Financial Modelling. Master of Science (MSc) Thesis. Department of Computer Science, University College, London. United Kingdom.

Deckmyn, P.H.A.M Davies, G.A & Bell, D.J (1994) Properties of Percolating Clusters on Finite Lattice Applied to model filtration processes. Elsevier Science Inc. Vol. 19

Dharmanna, L. Ramegowda, Manjunath, (2013). Fractal Dimension as Diagnostic Parameter to Detect Glaucoma. International Journal of Innovations in Engineering and Technology, Vol. 2, 63-69

Hackett, A. Gleeson, J. P. and Melnik, S. (2011) Site Percolation in Clustered Random Networks. International Journal of Complex Systems in Science. Vol. 1, pp. 25-32

Hamideh, K. (2011) Fractal Dimension Analysis of Grain Boundaries of 7XXX Aluminum Alloys and its Relationship to Fracture Toughness. Master of Science (MSc) thesis. Faculty of Aerospace Engineering, Delft University of Technology.

Iftekharaddin K.M, Jia W, Marsh A (2002) Fractal Analysis of Tumor in Brain Magnetic Response Images. Machine Vision and Applications Springer Verlag (2003) Vol 13: 352-362.

Micheal, J. L (2008) Methods in Percolation. Master of Philosophy (Ph. D) Thesis. Department of Mathematical physics, University of Canterbury.

Murthy K P N (2003) An Introduction to Monte Carlo Simulations in Statistical Physics. Indira Gandhi Centre for Atomic Research, Kalpakkam 603 102, Tamilnadu, India.

Reishofer, G. Koschutnig, K. Enzinger C. Ebner F. & Ahammer, H. (2012) Fractal Dimension and Vessel Complexity in Patients with Cerebral Arteriovenous Malformations: Francisco Jose Esteban, University of Jaen, Spain.

Salau, T. A. O. & Ajide, O. O (2012) Fractal Characterization of Dynamic System Sectional Images. Journal of Information Engineering and Applications, Vol. 2, 21-23.

Sankar, D. & Thomas, T. Fractal Features Based on Differential Box Counting Method for the Categorization of Digital Mammograms. International Journal of Computer Information Systems and Industrial Management Application (IJCISM) Vol. 2 pp. 011-019

Stauffer, D. & Aharony, A. (1992) Introduction to Percolation Theory, 2nd Edition. London: Taylor & Francis

Zhou H W and Xie H (2003) Direct Estimation of Fractal Dimension of a Fracture surface of Rock. Institute of Rocks Mechanics and Fractal, China University of Mining and Technology, P R China.