International Journal of Scientific & Engineering Research, Volume 6, Issue 6, June-2015 1
ISSN 2229-5518

Investigation of Variation of Cluster Size Distribution on Square Lattice for Various occupational Probabilities

Salau, T.A.O , Abayomi, K and Hashim, M.K

Abstract— This study investigated the variation in cluster sizes distribution on a square lattice using a Fortran program to populate square lattice haven specified the lattice size and corresponding occupational probability range. This program thereafter sorts the occupied lattice sites into good label using Hoshen and Kopelman algorithm. It then identifies all clusters present within the lattice, determine the cluster sizes distribution and group the cluster(s) according to size1, 2, 3, 4, 5, 10, 20, 30, 40, 50 and above respectively, With the entire procedure repeated for probability range starting from the initial (0.001) to the final (1.000) at a step size of 0.001 increments, tables and graphs results were then drawn, These results shows that the probability corresponding to the peak size distribution for all lattice studied increased toward threshold. Interestingly, the peak probability for size 50 and above for all studied cases suffered little absolute deviation relative to standard threshold value of 0.593 also the horizontal range of size distribution curve was noticed to reduce as size distribution increases, similarly an exponential relationship was noticed from the graph of horizontal range and corresponding size distribution for all studied cases which speculates that as x (size distribution) → ∞ that y (horizontal range) → 0 and vice versa.

Index Terms— Clusters, Hoshen and Kopelman algorithm, normalized graph, peak size distribution, percolation threshold, square lattice, step-size

---

1 Introduction

PERCOLATION as a mathematical theory was introduced by Broadbent and Hammersley [1], as a branch of probability that deals with properties of random media such as crystals and mazes, Percolation theory has also been used to described a stochastic way of modeling flow of fluid or gas through a porous medium of small channel which may or may not let the gas or fluid pass. In line with this Tammo et al, [2] used Percolation theory to explain water flow behavior in hydrophobic soil.

Eversince, this theory has been used to explain and model wide variety of phenomena ranging from economic, industrial, scientific and tremendous engineering applications, In view of dwindling financial market fluctuations, Damien [3] used percolation theory to study financial and market fluctuation using Count-Bouchard model.

In relation to medicine, Sharan and Shamir [4] used Percolation theory to analyze gene expression which requires clustering of genes into groups with similar patterns, a novel clustering algorithm called 'CLICK' was developed and applicable to gene expression analysis as well as to other biological applications. In line with this a simple equation has been given for optimal number of clusters and sample size per cluster, this equation was used for calculating the efficient sample sizes in cluster randomized trials with unknown intraclass correlation (ICC) and varying cluster sizes as opined by Gerard et al, [5]. In line with this, Yong and Lin, [6] used Percolation theory to model the spread of disease and propagation of forest fires.

Also, In relation to science and technology, Clifton and Lundquist [7] used a cluster algorithm "k-means cluster algorithm" to identifying clusters in wind data, haven studied the relationships between wind, turbine heights and climatic oscillations, the author thereby developed a method suited for predicting the impacts of climate change on wind resources, This method thereby allows objective identification of wind phenomena that may benefit the deployment of wind turbines.

Gerald and Eugene [8] used a bond percolating clusters at percolation threshold on a two-dimensional (2D) square lattices and three-dimensional cubic lattices was decomposed into singly connected ''links'' and multiply connected

- *Salau Tajudeen Abiola Ogunniyi is a senior lecturer in the Department of Mechanical Engineering, University of Ibadan, Nigeria.*
- *Abayomi Kayode is currently pursuing master degree program in mechanical engineering in University of Ibadan, Nigeria.*
  E-mail: abayomikayode66@yahoo.com
- *Hashim Mutiullah is currently pursuing master degree program in mechanical engineering in University of Ibadan, Nigeria.*

ISSN 2229-5518

‘‘blobs,’’ these blobs were then decomposed into a 3-block graph that cannot be separated into disconnected subgraphs using Monte Carlo simulations, the size distribution at which the blobs decomposed was determined as well as the fractal dimensions in relation to a concept ‘k-bone’ was also expressed. Again, Frey and Dueck [9] developed a clustering solutions algorithm termed “affinity propagation” (AP) which was an improvement over other older traditional data clustering procedures.

Michael and Köhn [10] developed a well-established heuristic for the p-median problem often obtains clustering solutions with lower error than the algorithm termed “affinity propagation” (AP) as well as other older traditional data clustering procedures.

Again, Hemnn [11] used percolation theory to model the geometrical growth of clusters (i.e. invasion percolation and diffusion in disordered media) which was used in the study of kinetic gelation processes.

A Monte Carlo simulation was used to investigate square lattice with distinct bond length in critical concentration, the results thereafter shows that the excition migration probability can actually provide a cluster generating function, as reported by Hosen et al [12].

These wide applications motivated this study with focus on the size distribution of ‘cluster’ present within a lattice. In engineering, model(s) and material(s) may not always be made of one (1) unique material(s), examples is a semiconductor material that is partly a conductor and partly an insulator, we can therefore arrange into a lattice as shown in Fig 1 so that cell(s) label ‘1’ would represent a property and cells labeled as zero (0) represents another property, it is important to note that this arrangement is a function of a chosen probability.

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

Fig. 1 A 5x5 lattice with occupational probability Pr=0.2

It is also however important to note that the position of each entry is independent of the position of its neighbor since the lattice is occupied randomly.

2. METHODOLOGY

A FORTRAN program was therefore developed to populate square lattice haven specified the lattice size and probability range [13][14][15][16]. This program thereafter sorts the occupied lattice sites into good label using Hoshen and Kopelman cluster finding algorithm [17], so that each cluster has its unique cluster label, the program thereafter determines the sizes of all clusters present within such lattice so that for each probability ten (10) trials were conducted (since the lattice is populated randomly using computer generated random numbers) so that an average result were taken for each occupational probability.

Cluster sizes were then grouped according to size1, 2, 3, 4, 5, 10, 20, 30, 40, 50 and above respectively. With the entire procedure repeated for all probability range starting from the initial (0.001) to the final (1.000) at a step size of 0.001 increments.

Tables of probability-corresponding size distribution was then drawn for all probability range from where the graph of size distribution against probability was draw using Microsoft excel also a table of probability-normalized size distribution was also drawn for all probability range from where graph of normalized size distribution against probability was also drawn using Microsoft excel.

Again, from the graph of size distribution against probability the horizontal range of all size distribution curve were taken at an arbitrary size distribution, in addition to this the peak size distribution for all size distribution curve were also recorded together with its corresponding probability, also the absolute deviation of this corresponding probability from threshold was also recorded. Hence a table of cluster size distribution, horizontal range, peak size distribution, its corresponding probability as well as its absolute deviation from threshold was then drawn.

From this table a graph of horizontal range against size distribution was plotted using Microsoft excel with the curve equation and curve fitness determined.

Similarly, from the graph of normalized size distribution against corresponding probability. a table of cluster size distribution, peak value of normalized size distribution curve, corresponding probability as well as its absolute deviation from threshold were taken and recorded.

These procedures were conducted for lattice sizes 200x200, 100x100, 50x50, 20x20 respectively.

3 RESULTS AND DISCUSSION

3.1 Results

TABLE 1
Selected sample results obtained for a 200x200 lattice at step size of 0.001, for occupational probability Pr (0.001 ≤ Pr ≤ 1.00)

	Size and above									
Probability	1	2	3	4	5	10	20	30	40	50
0.010	3932	73	3	0	0	0	0	0	0	0
0.011	4331	90	3	0	0	0	0	0	0	0
0.012	4782	108	5	1	0	0	0	0	0	0
0.013	5095	122	5	0	0	0	0	0	0	0
0.014	5378	148	6	1	1	0	0	0	0	0
0.015	5830	156	8	1	0	0	0	0	0	0
0.995	10	10	10	10	10	10	10	10	10	10
0.996	10	10	10	10	10	10	10	10	10	10
0.997	10	10	10	10	10	10	10	10	10	10
0.998	10	10	10	10	10	10	10	10	10	10
0.999	10	10	10	10	10	10	10	10	10	10
1.00	10	10	10	10	10	10	10	10	10	10

.

Summary of results obtained using 200x200 lattice size is shown in table 1 below.

From the results obtained in table 1 above, a graphs of cluster size distribution against probability was then drawn, this graph is shown in fig 2.

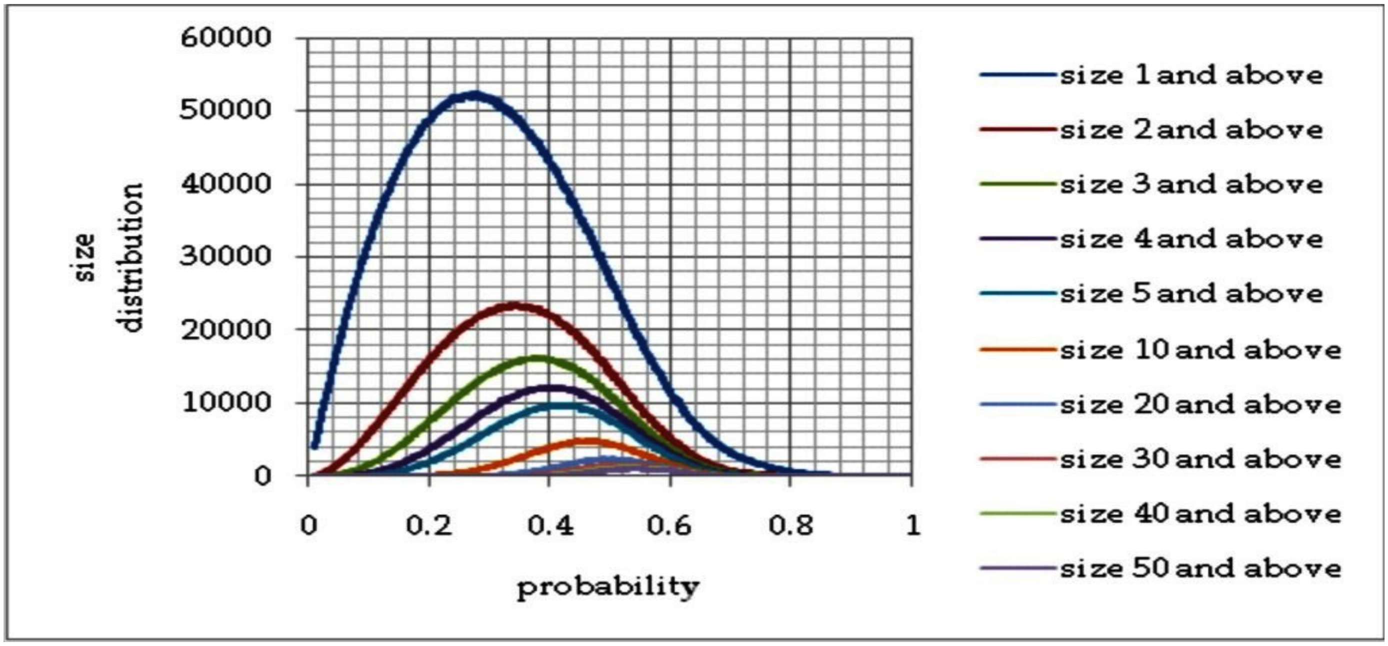

Fig 2: Graph of size distribution against probability for a 200x200 lattice.

Again, from the result summary in table 2, a table of probability-normalized size distribution was also drawn as shown in table 3.

Also, from table of result in table 1 above, a table of probability –normalized size distribution was drawn as

TABLE 2

Selected sample results for a probability-normalized size distribution obtained for a 200x200 lattice at a step size of 0.001 for occupational probability Pr (0.010 ≤ Pr ≤ 1.000)

	Size and above									
probability	1	2	3	4	5	10	20	30	40	50
0.01	0.075	0.003	0	0	0	0	0	0	0	0
0.011	0.083	0.004	0	0	0	0	0	0	0	0
0.012	0.091	0.005	0	0	0	0	0	0	0	0
0.013	0.097	0.005	0	0	0	0	0	0	0	0
0.014	0.103	0.006	0	0	0	0	0	0	0	0
0.015	0.111	0.007	0	0	0	0	0	0	0	0
0.995	0	0	0.001	0.001	0.001	0.002	0.004	0.006	0.008	0.01
0.996	0	0	0.001	0.001	0.001	0.002	0.004	0.006	0.008	0.01
0.997	0	0	0.001	0.001	0.001	0.002	0.004	0.006	0.008	0.01
0.998	0	0	0.001	0.001	0.001	0.002	0.004	0.006	0.008	0.01
0.999	0	0	0.001	0.001	0.001	0.002	0.004	0.006	0.008	0.01
1	0	0	0.001	0.001	0.001	0.002	0.004	0.006	0.008	0.01

shown in table 2.

ISSN 2229-5518

Hence, from table 2 a graph of normalized cluster size distribution against probability were plotted as shown in fig 3.

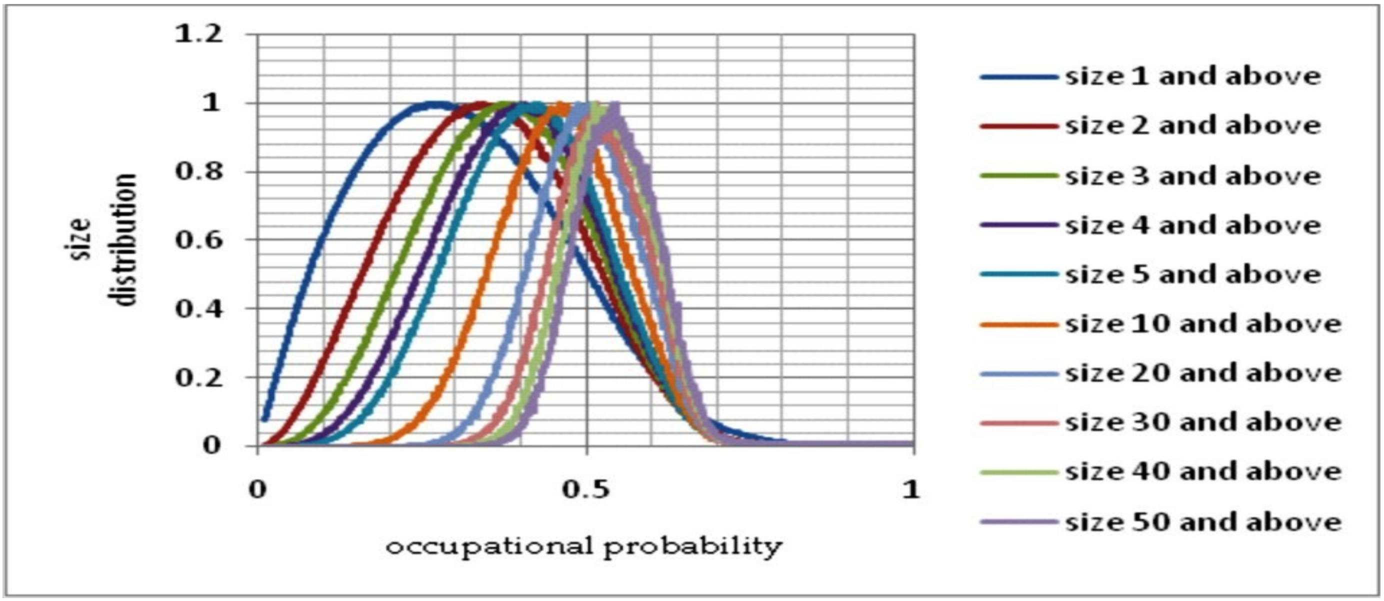

Fig 3: A graph of normalized cluster size distribution against probability

Also, from the graph of size distribution against probability (fig 2). The horizontal range at an arbitrary size distribution, its peak size distribution, its corresponding probability as well as its Absolute deviation from threshold was taken and recorded as shown in table 3.

Again from table 3 a graph of horizontal range at the said arbitrary size distribution of 4000 was plotted against the cluster size distribution as shown in fig 4.

TABLE 3

Table of cluster size distribution, horizontal range, peak size distribution, probability corresponding peak size distribution, absolute deviation from threshold for a 200x200 lattice.

Size and above	Horizontal range at an arbitrary size distribution of 4000	Peak Size distribution	Probability corresponding to the peak size distribution	Absolu deviati from thresho
1	0.66	52000	0.28	0.313
2	0.54	23000	0.34	0.253
3	0.45	16000	0.38	0.213
4	0.37	12000	0.40	0.193
5	0.31	9000	0.42	0.173
10	0.10	5000	0.44	0.153
20	-	3000	0.50	0.093
30	-	1500	0.51	0.083
40	-	1200	0.52	0.073
50	-	1000	0.54	0.053

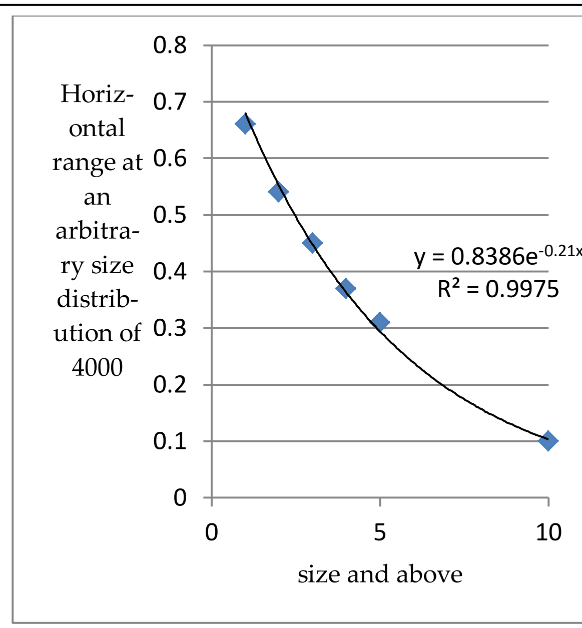

Fig.4. Graph of horizontal range at an arbitrary size distribution of 4000 against cluster size distribution.

ISSN 2229-5518

Also from the graph of normalized size distribution against probability (fig 3), a table of cluster size distribution as well as its absolute deviation from threshold was obtained and shown in table 4

TABLE 4
TABLE OF CLUSTER SIZE, PEAK NORMALIZED SIZE DISTRIBUTION AND PROBABILITY CORRESPONDING TO THE PEAK SIZE DISTRIBUTION

Size and above	Peak normalized Size distribution	Probability corresponding to the peak size distribution
1	1	0.28
2	1	0.34
3	1	0.38
4	1	0.40
5	1	0.42
10	1	0.44
20	1	0.50
30	1	0.51
40	1	0.52
50	1	0.54

Similarly, these entire procedures were also repeated for lattice sizes 100x100, 50x50 and 20x20 respectively, but due to space management only the graphs of horizontal range against cluster size distribution are shown in Fig 5, 6 and 7 for lattice sizes 100x100, 50x50, 20x20 respectively.

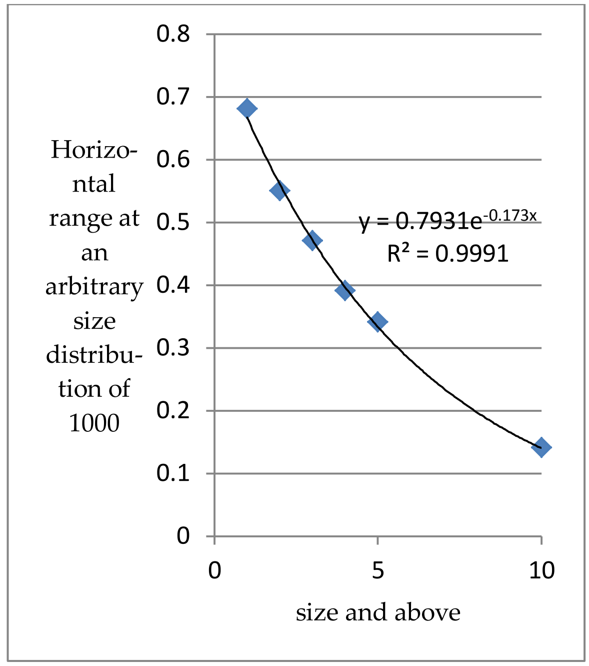

Fig.5. Graph of horizontal range at an arbitrary size distribution of 1000 against cluster size distribution for a 100x100 lattice size

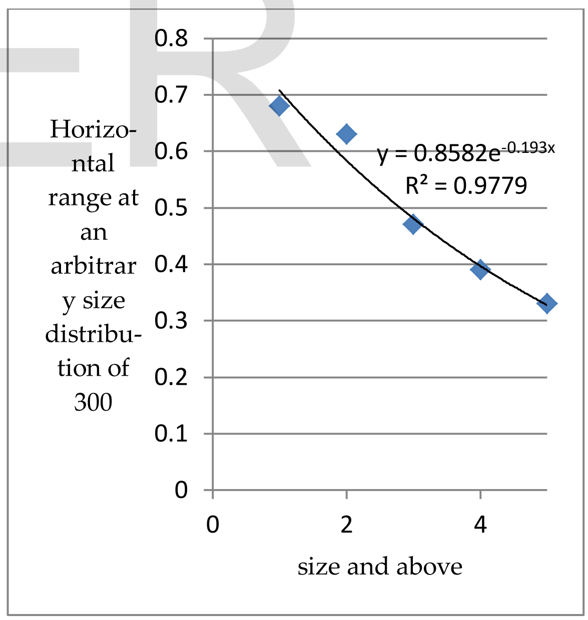

Fig.6. Graph of horizontal range at an arbitrary size distribution of 300 against cluster size distribution for a 50x50 lattice size

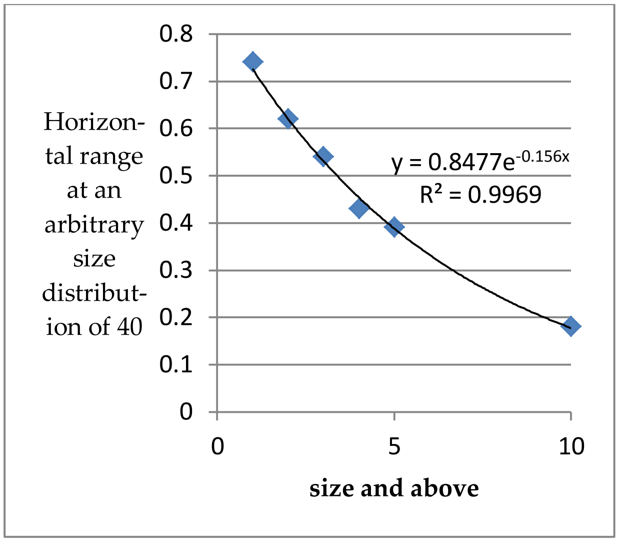

Fig.7. Graph of horizontal range at an arbitrary size distribution of 40 against cluster size distribution for a 20x20 lattice size

3.2 Discussion

At a probability of zero(0), No effect whatsoever would be expected for all lattice studied, but on increasing the occupational probability to say 0.01 (i.e. 1in every 1000 cells) the effect is almost unnoticeable on small lattice size, again on increasing the lattice size the transition becomes sharper and easily noticed.

On increasing the occupational probability to 0.593 (percolation threshold for a square lattice), a spanning cluster was noticed thereby validating percolation threshold for a square lattice as reported by [18].

From the graph of size distribution against probability (Fig 2) it was noticed that all sizes curves ascended at the unset, but on the other side begun to decline, thereby given a suspicion that all size curves has a peak value (though vary for different cluster sizes), this probability corresponding to the peak size distribution (fig2 and 3) was also noticed to increase with increasing size distribution, Also, the probability corresponding to the peak size distribution for tables 2 and 3 are the same which shows that the normalized graph is another representation of cluster size variation.

Again, from the graph of size distribution against occupational probability (fig 2) and graph of normalized size distribution against probability (fig 3), the horizontal range reduces as size increases, given the speculation that small cluster sizes distribution encompasses large cluster size distribution.

From the graphs of horizontal range at an arbitrary size distribution against size distribution (fig 5,6 and 7) an exponential curve equation was obtained with curve fitness in the neighborhood of 1, This was noticed for all cases studied.

4 CONCLUSION

For small lattice size the occupational probability must be large so that any effect or transition can be noticed easily. Probabilities corresponding to the peak size distribution for all lattice studied increased toward threshold. Interestingly, the peak probability for size 50 and above for all studied cases suffered absolute deviation of 0.007, 0.013, 0.033, 0.053 relative to standard threshold value of 0.593 for 20x20, 50x50, 100x100, 200x200 lattices respectively, Also size 1 and above curve has the largest horizontal range which reduces gradually until it reaches size 50 and above. In conclusion, from all graph of Horizontal range at an arbitrary size distribution against size and above an exponential relationship was observed speculating the prediction that as x (size distribution) tends to infinity and y (horizontal range) tends to zero and vice versa

REFERENCES

[1] S.R. Broadbent. and J.M, Hammersley, "Percolation processes" i and ii, *Proceedings of the Cambridge Philosophical Society* 53, 629–645. (1957).

[2] M. Tammo, G.H Allen, P. J-yies and P.E Robert, "Assessment of the application of percolation theory to a water repellant soil" Austrailia jornal of soil research 43, 357-360. (2006).

[3] D. Damien "On percolation theory for agent-based financial modeling" (september.2009).

[4] R. Sharan. and R. Shamir, "CLICK: A Clustering Algorithm with Applications to Gene Expression Analysis", (2000).

[5] J.P Gerard, van Breukelen*, J.J.M Math. Candel "Calculating sample sizes for cluster randomized trials: We can keep it simple and efficient" journal of clinical epidemiology 65, (2012).

[6]Z. Yong and G. Lin, "Journal Of Networks", Vol. 8, No. 8, Network Percolation Based On Complex Network. (August 2013)

[7] A. Clifton and J.K Lundquist, "Journal of Applied Meteorology and Climatology" Volume 51 Data Clustering Reveals Climate Impacts on Local Wind Phenomena. (2012).

[8] P. Gerald* and H. Eugene "Beyond blobs in percolation cluster structure: The distribution of 3-blocks at the percolation threshold", Physical Review E, Volume 65, 056126, (2002)

[9] B.J Frey and D. Dueck, Science Reports 315, p. 972), (16 February 2007)

[10] J. Michael Brusco1* and Hans-Friedrich Köhn "comment on "Clustering by Passing Messages Between Data Points" Downloaded from www.sciencemag.org on (February 11, 2015).

[11] H.J Hemnn *Phys. Rep.* 136 155-227 (1986).

[12] J. Hoshen , R. Kopelman, and E. Monberg, "Percolation and cluster distribution.ii. layers, variable-range interactions, and exciton cluster model", Journal of Statistical Physics 19, 3, 219{242. (1978)

[13] S.J Chapman, "Fortran 95/2003 for scientist and engineers" third edition'

[14] M. Metcalf and J. Reid "Fortran 90/95 explained" published by Oxford university press, (1999)

[15] W. Schink and G. Silverman, "Fortran 90 and engineering computation", Published by wiley, (1995).

[16] S.J Chapman, "Fortran 90/95 for scientist and engineers" published by McGraws-Hill, (1998).

[17] J. Hoshen and R. Kopelman, "Percolation and cluster distribution. i. cluster multiple labeling technique and critical concentration algorithm," Physical Review B 14, 8, 3438{3445, (1976).

[18] P.L Leath, "Cluster size and boundary distribution near percolation threshold," Phys. Rev. B 14, 11 (Dec 1976), 5046{5055, (December 1976).

ISSN 2229-5518