## ❖ Introduction :-

➢ Rather than just looking at the amount of space occupied by a single object, or how the object's space is configured, we will now concern ourselves with the overall arrangement of the objects in the space.

➢ Such an arrangements can be characterize by how many objects are there in a particular area, whether they are distributed evenly or in a group.

➢ We will look at the distance relationship between objects and how they are related to the overall size of the area.

➢ Objects like roads, fences, and houses can occur in a pattern. As we increase our geographical filter, we will see still more.

## ❖ Point, Area and Line Arrangement :-

➢ **Definition :-**
*"Spatial Arrangement is the placement, ordering, constriction, connectedness or dispersion of multiple objects."*

➢ **Why it is required?**
Spatial arrangement generally refers to the simple cartographic display of spatially distributed objects.

In such a case if we are using map for the communication purpose, then because of some reason spatially distorted features of the map would wrongly communicate.

Thus, for making the correct use of the map as a communication document we need to make an arrangement of such a feature.

## ❖ Point Pattern :-

➢ **The simplest measure of the point pattern is a matter of determine the _density of the distribution_.**

➢ This is done by dividing the number points by total area in which the points are existing.

➢ **For Example**: Population density, Housing density, Tree density are commonly used to measure the _compactness of point_.

➢ Generally four patterns are there for points:

1. Uniform pattern
2. Regular pattern
3. Random pattern
4. Clustered pattern

    ➢ Uniform Pattern :

       ▪ If the total numbers of points within the whole area are equally distributed in all the subareas of the whole area this type of point distribution is called as a Uniform Pattern.

    ➢ Regular Pattern :

       ▪ If a points occur on a grid separated by exactly the same distance throughout the entire area is called the Regular Pattern of point distribution.

➢ Random Pattern :

▪ Variable Distance between each pair of the points within the entire area is called the randomly distributed Point Patterns.

➢ Cluster Pattern :

▪ The points are grouped in a tight arrangement is called Cluster Pattern.

❖ **Quadrate Analysis :-**

▪ Uniform point patterns are defined based on the relationships among uniform subareas called ***Quadrates*** of the larger area.

▪ If each uniform quadrate contains the same number of point objects, we say the distribution is uniform for the overall study area. But it is not likely to found for biological phenomena because living organism migrates towards the specific locations that are rich in nutrient (પોષક).

▪ Thus we need to be able to test to see whether the distribution is uniform or not. The standard method for distribution testing is called ***Quadrate Analysis***.

▪ To determine how many numbers of points are distributed in each subarea of the total area, we divide the total number of points by total number of subareas, the number we get is called ***Expected Distribution***.

▪ Simple statistical test to applied to this set of data called **Chi-Square (X²) test**. We subtract the expected number of points (E) for each quadrate from the actually observed number (Q), the square of the result will then divided by the expected number (E) and take the summation of all sub areas.

▪ **Where :**

$$X^2 = \Sigma \frac{(Q - E)^2}{E}$$

o **E** = Expected Number of points for each subarea.
o **Q** = Actually Observed number of points for each subarea.

▪ **The result analysis of the Chi-Square test :**

o If the number we get is not substantially from what we would we expect, the distribution is uniform.
o *High Chi-Square value* indicates that our distribution is clustered.
o *Small Chi-Square value* indicates that the distribution is more evenly dispersed or uniform.
o *Intermediate Chi-Square value* indicates the distributed is more closely associated with the random process.
o Although quadrate analysis is help to decide whether the uniform distribution is exist in given situation, it can require more information as well.

▪ A special technique based on the quadrate is called ***Variance-Mean Ratio (VMR)***.

▪ It is an index that shows the relationships between the frequency of subarea and the average number of points in each subarea.

▪ It is calculated by the dividing the variance of the subarea point frequencies by the mean for each area.
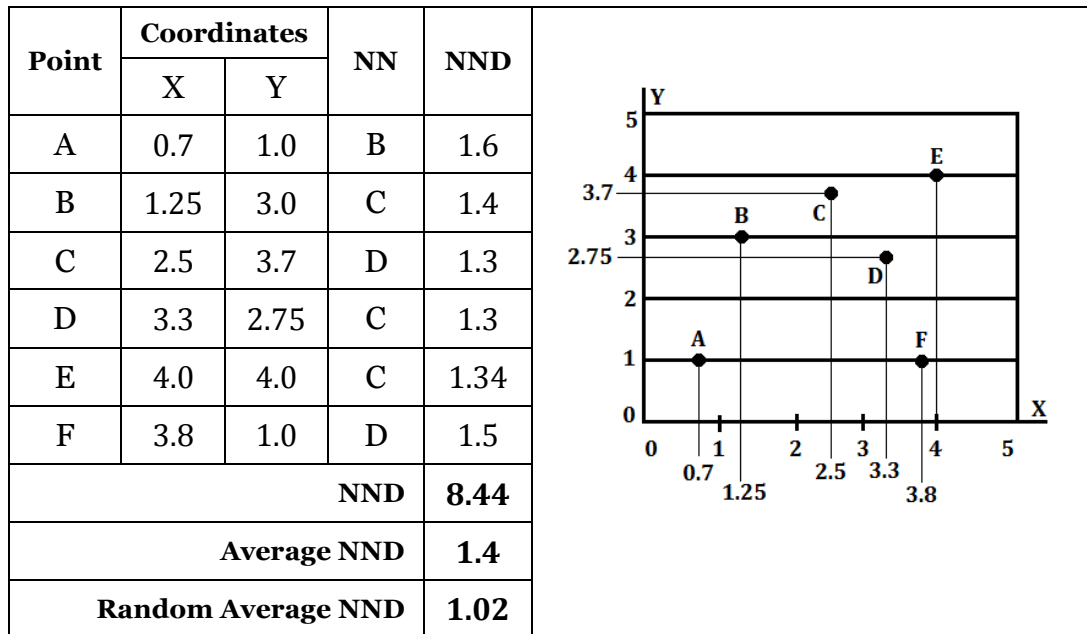
▪ **Where :**

$$VMR = \frac{VAR}{MEAN}$$

▪       VAR = Variance of the subarea point frequencies.
▪       MEAN = Mean for each area.

❖ **Nearest Neighbor Analysis :-**

- **Nearest Neighbor Analysis** is a common procedure to determining the distance of each point of its nearest neighbor and comparing that value to an average between-neighbor distance.
- The mean nearest neighbor distance provides a measure or index of the spacing between points in the distribution.
- It is useful because obviously point objects will conflict if they are too closer together.

| Point | Coordinates | | NN | NND |
|-------|-----|-----|-----|-----|
|  | X | Y |  |  |
| A | 0.7 | 1.0 | B | 1.6 |
| B | 1.25 | 3.0 | C | 1.4 |
| C | 2.5 | 3.7 | D | 1.3 |
| D | 3.3 | 2.75 | C | 1.3 |
| E | 4.0 | 4.0 | C | 1.34 |
| F | 3.8 | 1.0 | D | 1.5 |
| | | | NND | **8.44** |
| | | | Average NND | **1.4** |
| | | | Random Average NND | **1.02** |



- <u>6 points within 25 square unit</u>. **Density of Point = 6/25 = 0.24**
- An index of the random distribution, simply divided 1 by twice the square root of the density of point.
- **<u>Random Average NND</u>** $= \dfrac{1}{2*\sqrt{\text{Point Density}}} = \dfrac{1}{2*\sqrt{0.24}} = \dfrac{1}{0.98} = \mathbf{1.02}$

❖ **Thiessen Polygon :-**

- The points can be organized into a more regional context without relying on the use of another map layers for comparison by means of ***Thiessen Polygon***.
- It is also referred to as a **Dirichlet Diagrams** and **Voronoi Diagram**.
- Rather than simply looking at the arrangement of points, we can <u>grow a polygon around each point</u> to illustrate the "*region of influence*" <u>it might have no other polygon in the coverage</u>.
- The Thiessen Polygon is conceptually simple, but the process can become quite complicated as the number of point increase.
- If we have a number of point objects, such as town, we can imagine that *each point is surrounded by a single irregular polygon*.
- "*In most cases, each point in the coverage will have its own Thiessen Polygon, it is sometimes called **Proximal Region**.*"

- ▪ **How to create a Thiessen Polygon?**
  - ✓ We might conceptualize the Thiessen Polygons are derived by envisioning a bubble growing outward from each point, such that each of the interfaces between bubble from a straight line.
  - ✓ This straight line is oriented particular to a line that connect each pair of neighboring points.
- ▪ If you measure a distance between the two end points of the line, you will see that the distance is identical on either side of the line forming the interface.
- ▪ In other words, the edges of the polygons are formed by equidistance lines to the line connected each pain of the points.
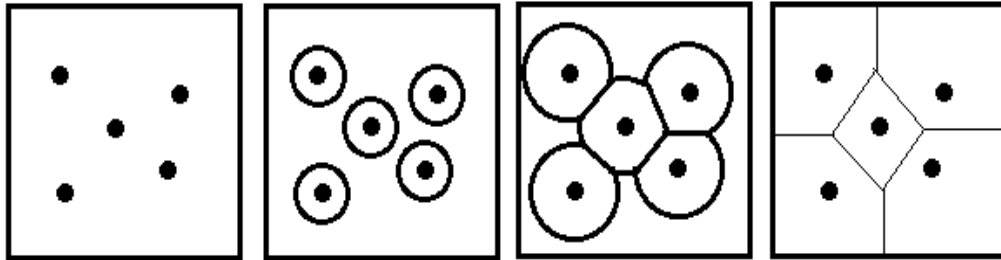


- ▪ **What we do with Thiessen Polygon?**

  - ✓ If we have a few scattered points of data but we want to characterized regions on the basis of these points, we use Thiessen Polygon to divide our space into a polygon form.

  - ✓ Most use of the a Thiessen Polygon involves determining the influence of a point data representing shopping center, industries, or other economically based activity.

## ❖ Measurement of Linear Pattern :-

- ➢ Lines occupy a space and occur in a particular configuration and patterns on the land. We see pattern of lines all the times, but often we miss them. *For Example*: Roads and Highway form a line with particular definable pattern.
- ➢ **Following are the methods for measuring the Linear Pattern.**
  1. Line Density.
  2. Nearest Neighbor and Line Intercepts.
  3. Directionality / Orientation of Linear Objects.
  4. Connectivity of Linear Objects.
  5. Gravity Model.
  6. Routing and Allocation.

### ☯ Line Densities :-

- ✤ Because of an added dimensionality of lines, especially length and orientation, linear analysis is somewhat more difficult than the study of points.
- ✤ Lines can occur frequently and infrequently with regards to the space they occupy.
- ✤ To measure a density value for point, which are consider to be dimensionless, we merely divided the number of points by the area.

- ✧ When we work with polygon, however, we accounted for the area occupied by each polygon by adding all the areas and dividing this sum by the area of the entire map.
- ✧ In the same way to measure the density of line we divide the sum of all the line length by the area of map coverage.

$$\textbf{Line Density} = \frac{\textbf{Sum of All Line's Length}}{\textbf{Area of Entire Map}} = \frac{\textbf{L}}{\textbf{A}}$$

## ☉ Nearest Neighbor and Line Intercepts :-

- ✧ <u>The Nearest Neighbor Statics can be used to characterized the spatial arrangement of Lines</u>. We can also determine the distribution of pair of lines using this technique.
- ✧ Because of heaving two dimensions our computation is more complex compare to the points.
- ✧ It might seem simple to adjust this additional dimension with the center of line and performing the nearest neighbor on that point only.
- ✧ Because the line objects are of variable lengths, this procedure is going to give us a true picture.

- ✧ **Steps for measuring the Nearest Neighbor of line objects :**

  - Select random points on each line on the map.
  - The distance between each point and its nearest neighbors is drawn to that line.
  - Then we measure the distance and calculate the mean nearest neighbor distance for all distance.

  

  Finding the Nearest Neighbor of the Line using a Randomly Selected Points along each line as the sample location

  - As with all nearest neighbor statics we need to be able to test this value against a random distribution.
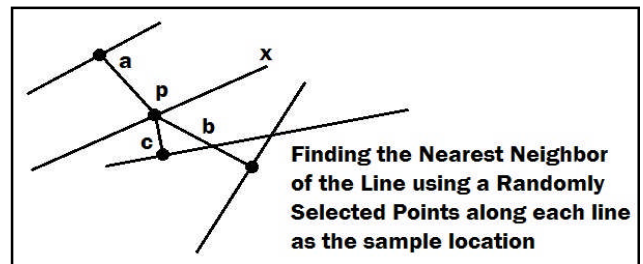
- ✧ This test will work for most line patterns, but ***it has some limitations***.
- ✧ *If lines in your coverage are very sinuous*, and especially if they change directions frequently, this approach is less than successful.
- ✧ **For the test to be useful**, <u>the lines should be at least 1.5 times the length of the mean distance between the lines</u>.
- ✧ <u>If the number of lines in the coverage is small</u>, the estimate of density used in nearest neighbor analysis should be **adjusted by a weighting factor of (n-1)/n**, where n is the number of lines.
- ✧ So rather than simply dividing the sum of the length by the area, we used the following formula : $\dfrac{\textbf{(n}-\textbf{1)L}}{\textbf{nA}}$

- ✧ **Where :**

  - **A** = Area

  - **L** = Sum of the Lengths of all the line objects

---

✧ **Line Intercept Methods :-**

▪ These are the alternative methods to analyzing the distribution of lines. One simple approach to convert two-dimensional pattern into a one-dimensional sequence by drawing a simple line across a map and notice where that line intersect the coverage line objects **(Draw a Random Sample Line) called _Random Walk_.**

▪ **At least two basic methods to produce a sample line :**

1. The first is to randomly select the pair of coordinate and connect them with lines.
2. A second method is to draw a radius at a randomly chosen angle and starting point, then measure a random distance from the center, finally constructing a perpendicular to a radius line.



**Random Walk Method
To Evaluate the Line Pattern**
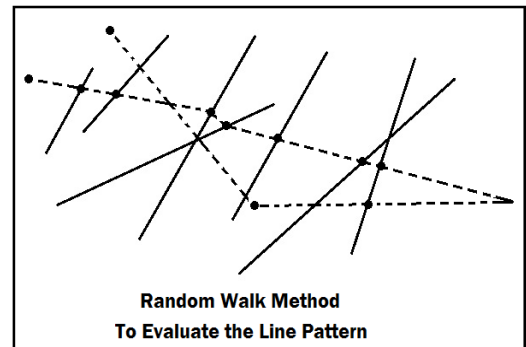
✧ Once a random sample line has been drawn over the coverage lines, the distribution of intersection intervals can be tested through the use of simple statistics like the _Runs Test_.
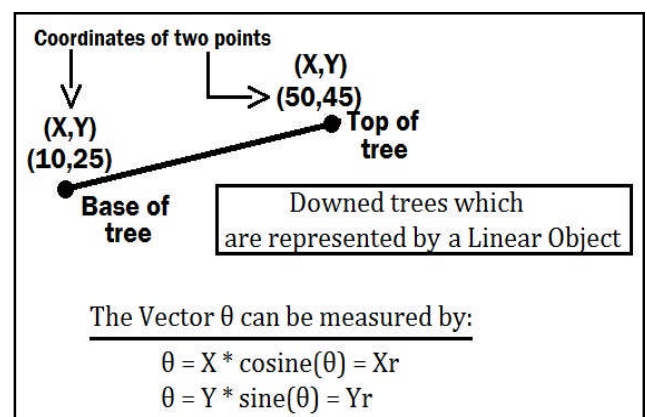
## ☯ Directionality / Orientation of Linear Objects :-

✧ In traditional statistical analysis of orientation, the data from map of linear objects are transferred to a circular graph called _Rose Diagram_.

✧ Which plots all linear features starting at the center of the circle and drawn each observation as a single line.

✧ Our first analytical technique is to figure out what the _Resultant Vector_ is.

✧ This is the procedure in which three people pulling an object in different direction and the vector gives us an idea where the object is likely to move.

✧ **For Example**, The large number of trees that were blow down by straight line winds associated with frontal activity.
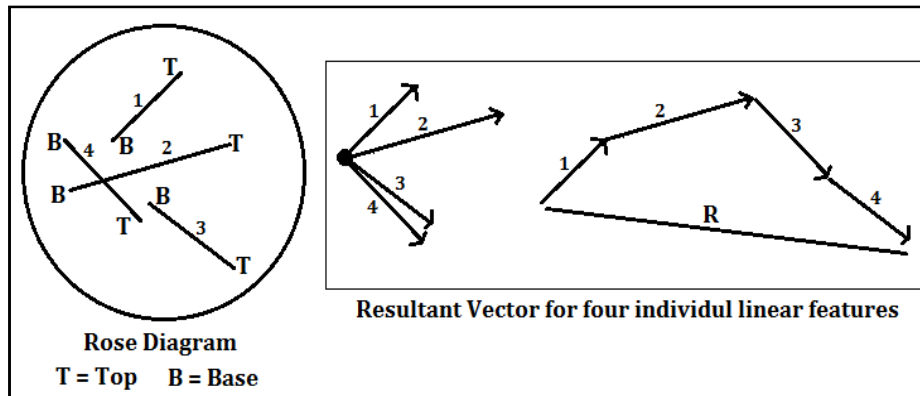
✧ **The top and bottom position** of each of the tree, will gives us both **orientation** and **direction** of every downed tree.

✧ The **vector** for each tree is defined by an angle **θ (theta)** measured from the base of the tree to its top position.



**Coordinates of two points**

(X,Y)
(50,45)

(X,Y)
(10,25)

**Top of tree**

**Base of tree**

Downed trees which are represented by a Linear Object

The Vector θ can be measured by:

$$\theta = X * cosine(\theta) = Xr$$
$$\theta = Y * sine(\theta) = Yr$$

✧ We multiply the X coordinate of each tree by the cosine of theta and each Y coordinate with the sin of theta.

✧ **To find the Resultant Vector**, We sum all these values for both coordinates, and the resultant vector values Xr and Yr show the dominant direction of the end points of all the trees in the blow down.

✧ **To determine Mean Direction θ base on Resultant Vector :**

✧ An average of the directions for all the downed trees is equivalent to the mean of any other set of data.

✧ The Formula for **MEAN** is : $\bar{\theta} = \tan^{-1}\left(\dfrac{Yr}{Xr}\right)$



**Rose Diagram**
**T = Top     B = Base**

**Resultant Vector for four individul linear features**

✧ We can standardize these values by dividing the coordinates by of each resultant vector by the number of line objects in the coverage, because the mean direction of our vectors depends not only on the desperation of the trees but also on the number of observations.

✧ This will allow us to compare two different areas.

✧ **For Example**, we could determine whether the winds were generally coming from the same direction.

✧ **To measure a Compactness of Individual Vectors :-**

✧ Resultant Vector is used to determine whether the individual lines are close together or widely spread.

✧ Following figure shows that 3 vectors are use to determine **Resultant Vector (R)**.



**Closely Spaced Vectors**

**Widely Dispersed Vectors**
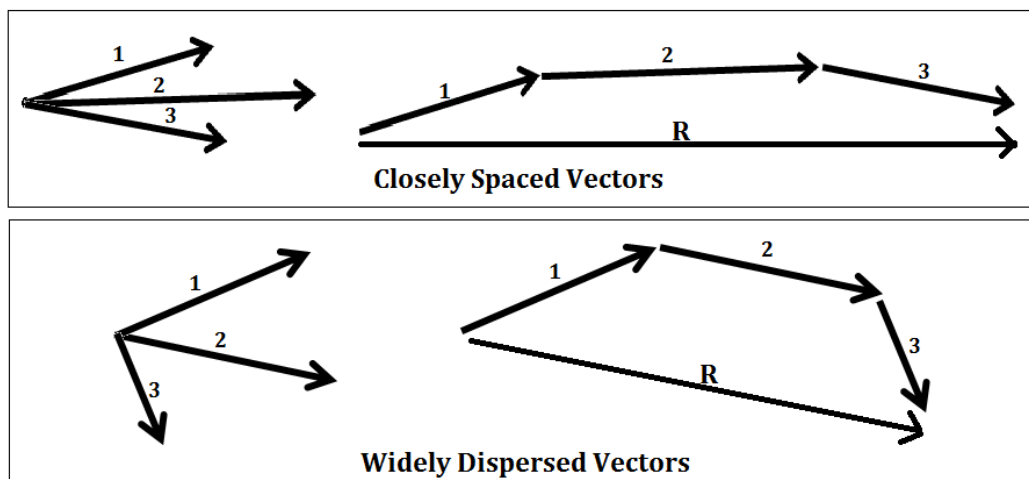
✧ When these vectors are closely spaced than the **resultant vector is too long**.

✧ If the **resultant vector is shorter** than the vectors are widely dispersed.

- We can determine the length of resultant vector R using Pythagorean theorem on our resultants (Xr and Yr) : **Resultant Length R** $= \sqrt{X^2 + Y^2}$
- **And Mean Resultant Length $\bar{R} = \dfrac{R}{n}$ $(0 \leq \bar{R} \leq 1)$**

  Where **n** = No. of Observations.
  - $\bar{R}$ **= o** than All the Vectors are there in <u>Totally Opposite Directions</u>.
  - $\bar{R}$ **= 1** than All the Vectors are there in <u>Uniform Directions</u>.
  - $\bar{R}$ **Near to o** than Vectors are <u>Widely Dispersed</u>.
  - $\bar{R}$ **Near to 1** than there are <u>Closely Spaced Vectors</u>.
- **The Problem related to the orientation :-**
  - **Problem :**
  - The orientation of any linear objects gives two possible directions. Unlike trees, some features can be expressed in either of two directions.
  - **For Example**, two researchers are standing in opposite directions: One researcher says that most of the fencerows are oriented to north and other one is saying that most of the fencerows are oriented to south. <u>Both are true</u>.
  - **Solution :-**
  - Double the angle no matter which direction was originally used to record the data. Than calculate the resultant length **(R)** and mean resultant length **($\bar{R}$).**
  - To get our true value for these measurement, we all need to do is Divide by 2.

### ❷ Connectivity of Linear Objects :-

- Connectivity of linear objects means their ***ability to form networks***.
- Networks occurs in many different ways :
  - Natural and
  - Anthropogenic.
- Examples of Networks are:
  - Rail and Roads (Anthropogenic).
  - Telephone Lines (Anthropogenic).
  - Rivers and Streams (Natural) etc...
- With aspect of density and orientation (directionality), we need to able to analyze the actual connection made by these features and the amount of connectivity provided from place to place.
- *Connectivity is measure of the complexity of a network*.
- Several devices are used to calculate this value, the two most common once are :
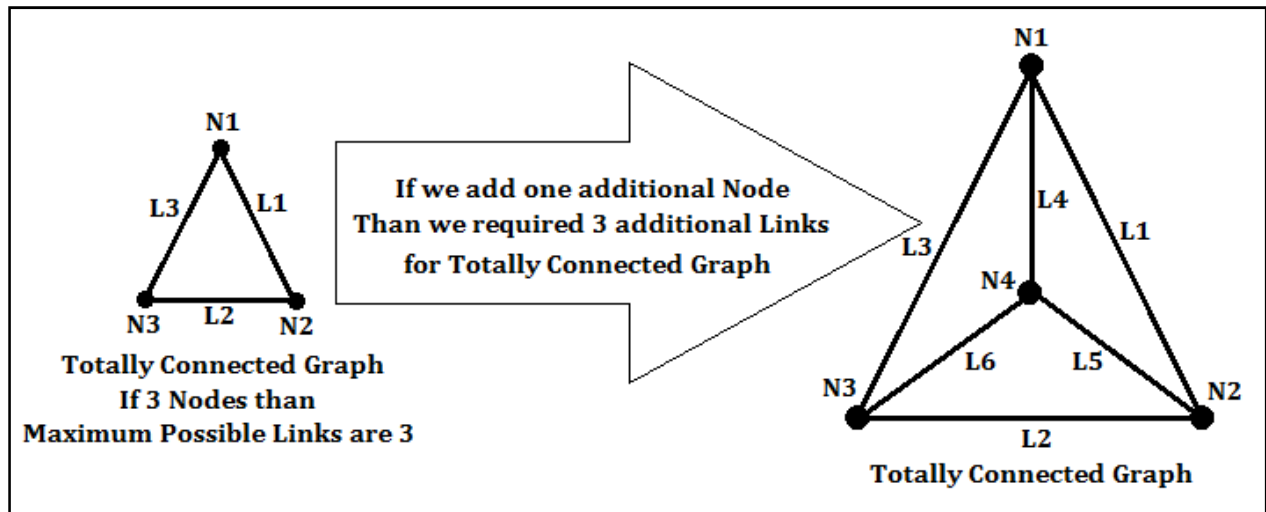  - **Gamma Index** and
  - **Alpha Index**.

## 1. Gamma Index :-

- Gamma Index (γ) compares the number of links (**L**) in given network to the number of possible links (**L$_{max}$**) between nodes.
- To calculate this we simply produce a **ration of the L and L$_{max}$.**
- First we count **the number of links that actually exist** in coverage (**L**).

↬ Once we have the actual number of links, **we need to determine the number of possible links in the coverage**.

↬ **For Example** : only three nodes are present, only three links are possible. But if we add another node, we see that three additional links are possible.
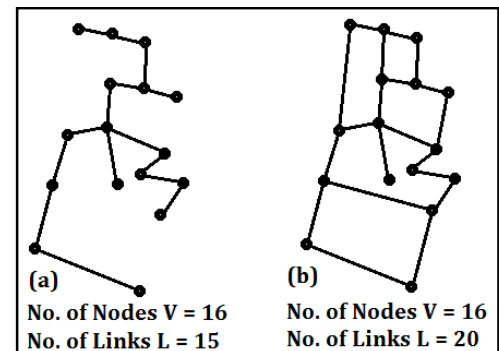


**Totally Connected Graph**
**If 3 Nodes than**
**Maximum Possible Links are 3**

If we add one additional Node
Than we required 3 additional Links
for Totally Connected Graph

**Totally Connected Graph**

↬ So if we assume, that no new interactions are formed, the maximum number of possible links is increased by three times.

↬ In other words, the maximum number of links, $L_{max}$ is always **3(V-2)**, where V is the number of nodes.

$$\gamma = \frac{L}{L_{max}} = \frac{L}{3(V-2)}, (0 \leq \gamma \leq 1)$$

↬ **Examples :-**

(a) $\gamma = \frac{L}{L_{max}} = \frac{L}{3(V-2)} = \frac{15}{3(16-2)} = 0.36$

(b) $\gamma = \frac{L}{L_{max}} = \frac{L}{3(V-2)} = \frac{20}{3(16-2)} = 0.48$



(a)
No. of Nodes V = 16
No. of Links L = 15

(b)
No. of Nodes V = 16
No. of Links L = 20

↬ Network (a) is 36% connected, whereas the Network (b) is 48% connected.

↬ If number of actual link (L) is equal to Number of Possible Links ($L_{max}$) than Gamma Index ($\gamma$) = 1 or 100% that means the network is fully connected.

↬ **L = $L_{max}$** than $\gamma$ = 1 = 100% = **Fully Connected Network**

## 2. Alpha Index :-

↬ The index design to measure the circuitry (the degree to which nodes are connected by circuit of alternative routes) is called Alpha Index ($\alpha$).

↬ Alpha Index is a ratio of the existing number of circuits to the maximum possible number of circuits.

↬ Network with no circuit present contains one link fewer than the number of nodes: means **network with no circuit heaving total L = (V-1) Links**.

✧ Therefore the **number of circuit present in the network**: (number of links present – number of links in our minimally connected network) can be given by **(L - V) + 1**

✧ Maximum number of circuits = number of possible links – links required for minimally connected graph [3(V-2) – (V-1) = **2V-5** ]

$$\alpha = \frac{\text{Actual No.of Circuits}}{\text{Maximum No.of Circuits}} = \frac{(L-V)+1}{2V-5}$$

✧ **Examples :-** (Consider the above given Networks (a) and (b))

(a)→ $\alpha = \frac{(L-V)+1}{2V-5} = \frac{(15-16)+1}{(2*16)-5} = \frac{0}{27} = \mathbf{0}$

(b)→ $\alpha = \frac{(L-V)+1}{2V-5} = \frac{(20-16)+1}{(2*16)-5} = \frac{5}{27} = \mathbf{0.19}$

✧ There is no circuit are present in Network (a) and 19% circuits out of 100% are present in network (b).

✧ **Use of Alpha & Gamma Indices :**
   ▪ However the alpha and gamma indices give us a different look at the overall network patterns, they might more appropriately combine to provide an overall amount of the ***network complexity***.

✧ To perform alpha and gamma indices in GIS requires that the GIS be a Vector system.

---

☞ **Gravity Model :-**

✧ This allows us to concentrate on the degree to which the links connected these nodes or the number of ***circuits and alternate paths available***.

✧ Regarding two examples of contrasting nodes, you should note that we have defined a magnitude for each that we can use to separate them out in a hierarchical fashion. That means to give a weight according to the use.

✧ The magnitude of the alteration can be thought of much as astronomers envision the gravitational attraction between two outer space bodies.

✧ *"Converting the concept of gravitational attraction to use in a two dimensional setting provides the measure of the interaction between two nodes in GIS coverage that we call the **Gravity Model**."*

$$\textbf{Gravity Model} \rightarrow \textbf{Lij} = \textbf{K}\frac{\textbf{Pi Pj}}{\textbf{d}^2}$$

**Where,**
   $L_{ij}$ is the interaction between node i and j.
   $P_i$ is the magnitude of node i
   $P_j$ is the magnitude of node j
   $d$ is the distance between the two nodes
   $K$ is the constants relates the equation to the types of objects being studied
   (For Example Population Size, Animals etc...)

✧ As with gravity the larger the magnitude of the node, the grater the interaction between them likely to be.

---

### ☯ **Routing and Allocation :-**

✧ Among the most useful applications of network in GIS are related to task of routing and allocation.

### ✧ **Routing :-**

- *"Routing involves finding the shortest path between any two nodes in a network."*
- Each link in the network can also be <u>assigned an impendence value</u>, much like a friction surface but imposed only on the line itself. This value might be related to the speed limit along a street.
- *<u>Based on the combination of the</u> **calculated distance** <u>and the</u> **impedance factor**<u>, the most efficient route can be found, rather than just a shortest path</u>.*
- Although routing can be done in the raster, Because of the close relationship of the graph theory topology and the topological vector data model, it is much easier if performed in a vector system.
- You should also be aware that given the many possible routes, especially where circuits are involved, there are likely to be a number of possible ways for finding your route.

### ✧ **Allocation :-**

- *"Allocation is the process that can be used to define, for example, the location of the market, the areal extent of a water treatment center, or the boundaries of a series of fire station service areas."*
- A network structure within a vector GIS is most often used.
- The idea here is that the <u>*capacity of a given service is distributed through the network*</u>.