Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 4s (2024)

Connected Graph with Bacterial Graphs and Network Distance

T. Selvaraj[1], Dr. S. Subramanian[2]

[1]Research Scholar Enroll No:16R1612458
[2]Professor& Head, Dept. Of Mathematics,
[1,2]PRIST Deemed to be University, Thanjavur-613403.
Email: jobu.selvaraj@gmail.com

Article History:

Received: *19-04-2024*

Revised: *03-06-2024*

Accepted: *22-06-2024*

Abstract

Network theory is going to analyse the set of techniques. But in a complex system network in a network of complex it has techniques to analyse the structure in a system of interacting agents. The graph theoretic representation means the system is made to apply network. The problem may be converted into graph by two components namely nodes and edges nodes are called as entities and interactions between edges are called as edges.

Keywords: Connected Graph, Network, bacterial, Distance, Components.

Tracking network dynamics using graph distances

Here in a graph a node is assigned to each entity (or) player and interactions are modelled by edges. Graphs have endless encapsulating structural information in data sets. It may become paradigm of indispensable here the education about every separate node marked as interactable (or) irrelevant. Intended recording each individual neuron's activity. The neurons are modelled by brain activity from brain data.

Raw data can be transferred as a group of graphs. input is taken as treatable. In the identify notes some set of aligned notes are very important, ever though distances constitutes a mile stone to any analysis of statistics simultaneously comparisons, variance (or) cluster (or) even the definition of median (or) mean. Once the distance has been selected, the crucial is the analysis graph data. The distance between two unlabelled graph are very important and it must be well studied. But our focus is different nodes are marked as identity. The relevant information must discard by permutation invariant distances. the node's leverage function is defined by nodes of the intensity of key node locations. with the distance sensitivity but of the self-graph, properties of distanced by exhibit distinct capture various types of structure Shifts correct scale for comparing graph is needed to identify decisive parameter.

If we want to do the local structural changes at an atomic level structural similarities must be studied the node communities of the level notation of the   heart is the scale which is used to find out different distance similarities between graph. The distances may be local, global and meso scale.
Template are used to select the data

Local - structural change

Global – spectral distance change

There are two categories to capture in between the two meso scale is in between the local and global a fresh group of graph distances depends on spectral heat kernels and argue that should be optimum and must combine both local and global the different performance is around on both synthetic, fMRI data and microbiome are controlled experiments

Vertice = ~

Edges = E

N=|~| count of nodes

|E|= count of edges

j – i neighbour nodes

D = Degree of matrix

A_{ij} = { 1 if i~j and 0 otherwise

D = diagram (d_i) = I =1,….,N

Such that,

$$d_i = \sum_{j=1}^{N} A_{ij}.$$

A matric in symmetric

$A^T = A$

L = – A+ D

(Laplacian's)

Laplacian is symmetric and its true eign value decomposition

$L = U \wedge U^T$

U = Unitary Matrix

^ = diag(λ_i) is the diagonal matrix of the eign values

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1}.$$

Quantifying structural distances via local changes

Distances between graph are only mutually exclusive (or) structural version. Spectral distances in local changes the graph structure at each node is different but it induces local changes.

Example Different bonds in different molecules are having different properties. But in brain network these distances can be noted as (ROI) Individual region of Interest. The other is identified by graph structure change in the Laplacian eign values (or) its adjacent matrix, distance marks the graph properties how nodes are functioning with nodes (or) globally organized and interact. In brain networks change in connectivity is assured globally with spectral distance.

The hamming distance

The count of edge insertions and deletions necessary to transfer in between two graphs. let Ġ and G are 2 graphs on nodes N. There adjacent metrices are A and Aˋ Hanning distance is defined by

$$d_H(G, \tilde{G}) = \sum_{i,j} \frac{|A_{ij} - \tilde{A}_{ij}|}{N(N-1)} = \frac{1}{N(N-1)} \|A - \tilde{A}\|_1.$$

bounded Distance between 1 and 0 in over all graph of sizes 'N'

Analysis of frame work in General

The frame work and graph are related one. pairwire dissimilarities among the graph in the MatricH is n by n dimensional matrix H

Where,

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 4s (2024)

Dimensional matrix, $H_{ij} = d(G_i, G_j)$

N = number of graphs

N =21 micro biome

N =29 fMRI data set

H can be analysed in various ways

Every distance is used for the analysis of (c) the graph variability from time to time is mentioned by various distance between next to next graphs

H between graphs [fig 1(A) ,E] and their low dimensional projections. It shows how much relationship it captures between connectivity and count of years only that subject under time passed cocacine relativity.

Application

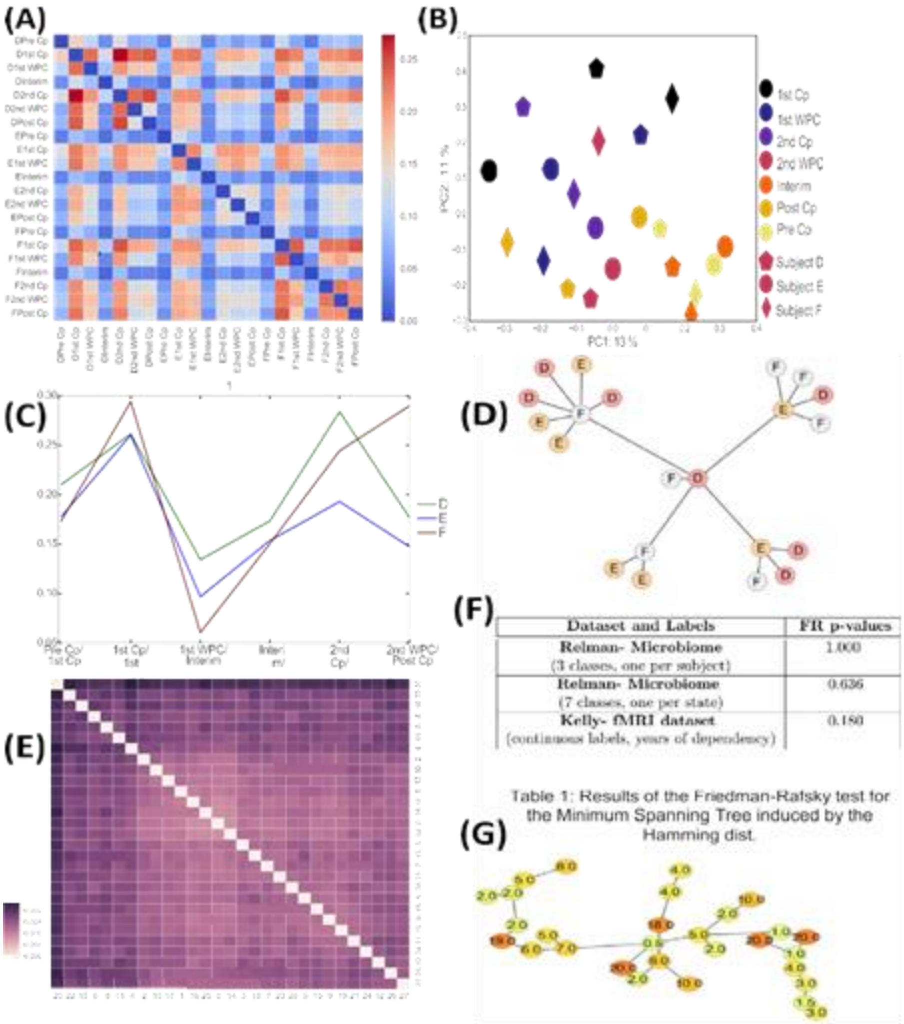


Fig [1] mention the result of microbiome on both bacterial using hamming distance as well as fMRI data.

Result

The fig (1) shows the bacterial community graph in which hamming distance is completed throughout the treatment course.

In fig 1[c] the distance between 2 consecutive graphs just often antibiotic courses c

Fig [1] B shows the treatment gradient

A - hamming distance between the bacterial graphs tops rows and brain graph located at bottom

bacterial graph – A

B – MDS projections of the first two principal component

Colour – phaser

Shape – different subject

In the graph c plot the next to next distance in between bacterial graph [c]

G – distance between minimum spanning tree G hamming distance

F – friedman – refsky test for significance for the diff data set

E – cluster map of the fMRI graph

G – minimum spanning tree between brain connections induced G

The Jaccard distance

It is union graph with respect to volume in normalization

$$d_{\text{Jaccard}}(G,\tilde{G}) = \frac{|G \cup \tilde{G}| - |G \cap \tilde{G}|}{|G \cup \tilde{G}|} = \frac{\sum_{i,j} |A_{ij} - \tilde{A}_{ij}|}{\sum_{i,j} \max(A_{i,j}, \tilde{A}_{ij})}| = \frac{\|A - \tilde{A}\|_1}{\|A + \tilde{A}\|_*},$$

where ||.|| * mentions the Matrix of nuclear form. The above notation denotes the sharp distance between the graphs.

According to steinhus transform

$$\delta(x,y) = \frac{2d(x,y)}{d(x,c)+d(y,c)+d(x,y)}$$

Produces c matric

D = hamming distance

C = empty graph

$$\delta(G,\tilde{G}) = \frac{2\|A - \tilde{A}\|_1}{\|A\|_1 + \|\tilde{A}\|_1 + \|A - \tilde{A}\|_1} = \frac{2(|G \cup \tilde{G}| - |G \cap \tilde{G}|)}{2|G \cup \tilde{G}|} = d_{\text{Jaccard}}(G,\tilde{G}).$$

A - distances between bacterial graphs and Kendall-correlation by Jaccard

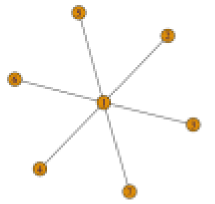
B - consecutive distances in between two bacterial group

C – bacterial’s MDS projection

D - Minimum Spanning Tree between bacterial graphs

Edge-importance

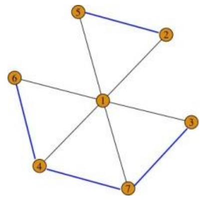
disconnected components should be penalized more.



Initial graph G_1

Edge-submodularity

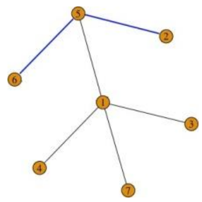
In a same nodes edge get charged



Perturbation of G_1 four edge insertions

Weight awareness

It denotes weight of modified



Perturbation of G_1:2 edge insertions,2 deletions

Focus awareness

random changes in not important not even happened Jaccard and humming distance treat all edges uniformly.

Comparing graph structures

Spectral distances are global = eign values of either (or) adjacency matrix

$$\mathbf{L = D - A}$$

L = Laplacian
D =diagonal matrix
A = adjacency matrix

Normalized Laplacian,

$$\dot{\mathbf{L}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\,\mathbf{D}^{-1/2}$$

L_p distances on the eigen values

$$d(G,\tilde{G})^p = \sum_{i=0}^{N-1} |\lambda_i - \tilde{\lambda}_i|^p .$$

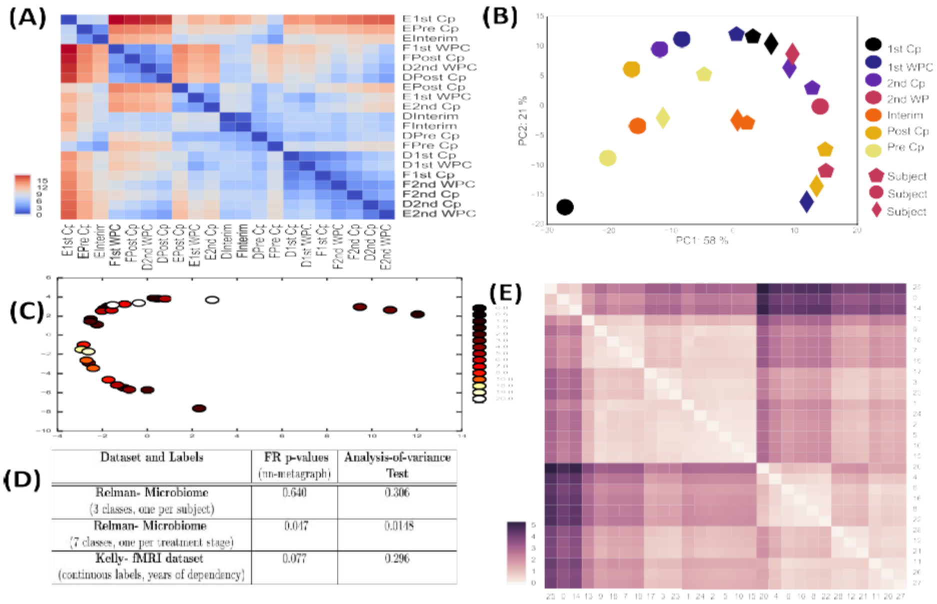
Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 4s (2024)

$$\lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1}$$

$$d_f(G, \tilde{G})^p = \sum_{i=0}^{N-1} |f(\lambda_i + \varepsilon_i) - f(\lambda_i)|^p \approx \sum_{i=0}^{N-1} |f'(\lambda_i)|^p |\varepsilon_i|^p$$

……Eqn 3.1

Applications



l_2 = distance by two different functions such as combinatorial Laplacian eign spectrum eqn (3.1)

$f(\lambda) = e^{-0.1\lambda}$ top row Microbiome

$f(\lambda) = e^{-1.2\lambda}$ bottom row fMRI

p = 0.025 confirmed the MDS projections

l_2 = eign spectra using two function of spectral distances as in eq no 3.1

top row = Microbiome

bottom row = fMRI

E = brain connectomes

C = first two principal axes with graphs

D = low-pass filter $f(\lambda) = e^{-1.2\lambda}$

B = MDS projection of the bacterial

A = distances between bacterial graphs

Spanning tree dissimilarities

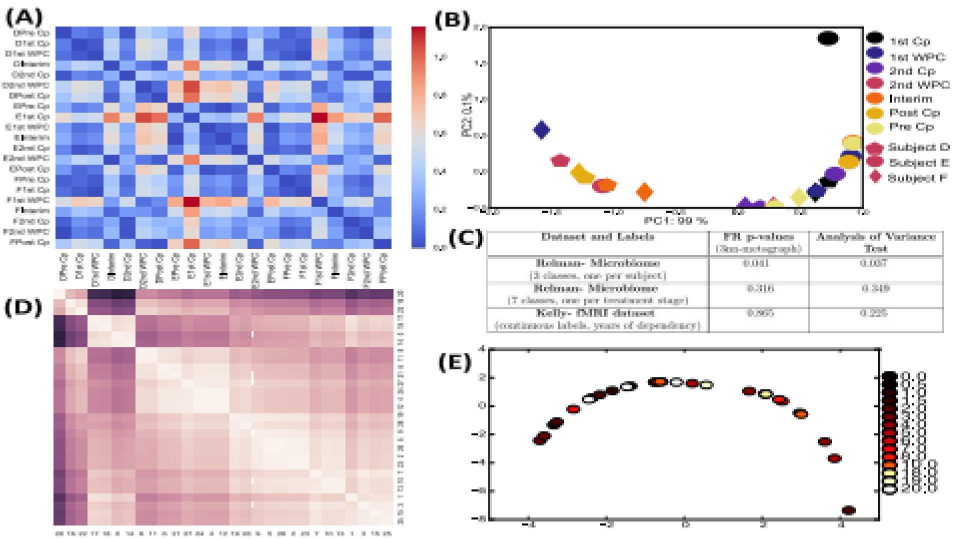
$$\mathcal{T}_G = N_{\text{Spanning tree of } G} = \frac{1}{N} \prod_{i=1}^{N-1} \lambda_i$$

G and $\dot{G}$ - graphs

L = D - A

$d_{ST}(G, \dot{G}) = |\log(T_G) - \log(T_{\dot{G}})|$

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 4s (2024)



taking network changes

A = bacterial graphs based on Kendall-correlation

B = bacterial graphs based on MDS projection on the first two principal axes

Colour – phaser

Shape – different subject

C = metagraph

D = fMRI data for Clustermap

E = brain connectomes MDS projections on the 1st two principal components(E)

$$\forall i, \qquad \tilde{\lambda}_i = \lambda_i + \varepsilon_i.$$

$$\tilde{T}G = N\text{Spanning tree of } \tilde{G}$$

$$= \frac{1}{N} \prod_{i=1}^{N-1} \tilde{\lambda}_i$$

$$= \mathcal{T}_G \times \left[1 + \sum_{i=1}^{N-1} \frac{\varepsilon_i}{\lambda_i} + \sum_{i,j=1}^{N-1} \frac{\varepsilon_i \varepsilon_j}{\lambda_i \lambda_j} + \cdots \right]. \quad \textbf{........Eqn (3.3)}$$

Combining (3.3) and (3.2) yields

$$d_{\mathrm{ST}}(G, \tilde{G}) = \left| \log\left(1 + \sum_{i=1}^{N-1} \frac{\varepsilon_i}{\lambda_i} + \sum_{i,j=1}^{N-1} \frac{\varepsilon_i \varepsilon_j}{\lambda_i \lambda_j} + \cdots \right) \right|.$$

General framework

For Gaussiankernel, distribution

$$\rho_G(x) = \frac{1}{n} \sum_{i=0}^{n-1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\lambda_i)^2}{2\sigma^2}}.$$

Pseudo graph,

$$d(G, \tilde{G}) = \int \left| \rho_G(x) - \rho_{G'}(x) \right| dx.$$

Definition of the IM distance

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 4s (2024)

$$\frac{\partial^2 x_i}{\partial t^2} + \sum_{j \neq i} A_{ij}(x_i - x_j) = 0.$$

The eign values of the squares of the vibrational frequencies which is inter- preted by Laplacian matrix of the network as ω_i.

$$\lambda_i = \omega^2$$

$$\lambda_0 = \omega_0 = 0$$

$$\rho(\omega, \gamma) = K \sum_{i=1}^{N-1} \frac{\gamma}{\gamma^2 + (\omega - \omega_i)^2}$$

γ = common parameter

K = normalization constant defined such that

` The spectral distance between 2 graphs A & B is defined by

$$\varepsilon_\gamma(A, B) = \sqrt{\int_0^\infty [\rho_A(\omega, \gamma) - \rho_B(\omega, \gamma)]^2 \, d\omega}.$$

Jurman,

$$\varepsilon_{\bar{\gamma}}(\mathcal{E}_N, \mathcal{F}_N) = 1.$$

$$\{A, B\} = \{\mathcal{E}_N, \mathcal{F}_N\}$$

E_N = empty graph on N Nodes

F_N = Complete graph on N Nodes

polynomial approach

$$P(x) = x + \frac{1}{(N-1)^\alpha} x^2 + \cdots + \frac{1}{(N-1)^{\alpha(K-1)}} x^K$$

$$P(A) = QW Q^T$$

Where,

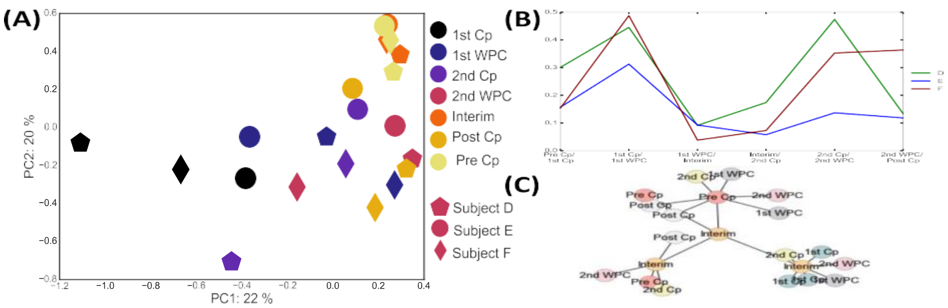
$$W = \Lambda_A + \frac{1}{(N-1)^\alpha} \Lambda_A^2 + \cdots + \frac{1}{(N-1)^{\alpha(K-1)}} \Lambda_A^K.$$

The distance between 2 graphs G_1 & G_2

$$d_{\text{poll}}(G_1, G_2) = \frac{1}{N^2} \| P(A_1) - P(A_2) \|_{2,2}.$$

Where,

$$\| M_1 - M_2 \|_{2,2} = \left(\sum_{i,j} |M_{i,j}^1 - M_{i,j}^2|^2 \right)^{\frac{1}{2}}$$



dissimilarity of the microbiome bacterial graphs which is polynomial

3 = K

0.9 = α

A = Heatmap of the corresponding similarity

B = MDS projection

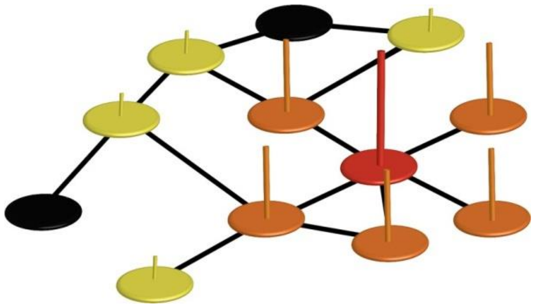
C = consecutive distances between graphs

The common network generally compare the dissimilarities between graphs next to next points (G_t and G_{t+1})

$$d_{\text{centrality}}(G_t, G_{t+1}) = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} (s_{ij}^{(t+1)} - s_{ij}^{(t)})^p \right)^{1/p}$$

G_t , G_{t+1} = Consecutive points

S_{ij}^{t} = affinity of interaction between i with j in graph G_t



C. Donnat and S. Holmes

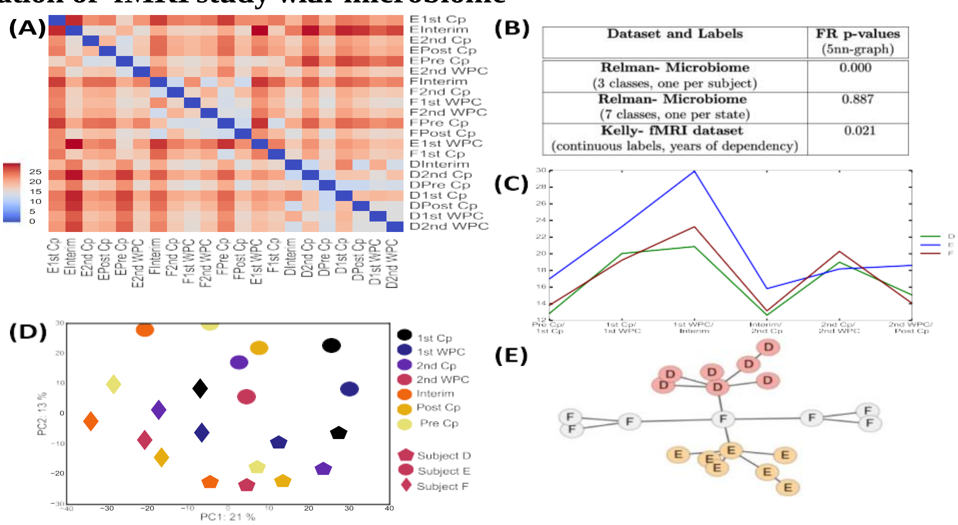
d(G_t , G_{t+1})

$$d(G_t, G_{t+1}) = \frac{1}{N} \sum_{a \in \mathcal{V}} \| r_a^{(t)} - r_a^{(t+1)} \|_2$$

$$= \frac{1}{N} \sum_{a \in \mathcal{V}} \delta_a \Delta^T \Delta \delta_a$$

$$= \frac{1}{N} \mathrm{Tr}[\Delta^T \Delta]$$

Application of fMRI study with microbiome



results of heat based distance with τ =1.2

P value is below 0.05 threshold, 5-nn metagraph is obtained

Fig 9(B), (C), (D), (E) are the some one of the above the Annova gives the result a P value below10^{-4}

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 4s (2024)

Synthetic experiments

the graph topology

the Erdős–Rényi model = 81nodes probability p = 0.1

Preferential Attachment (PA) = N = 81 nodes α = 1, a Stochastic Block Model with 3 sampled equally communities

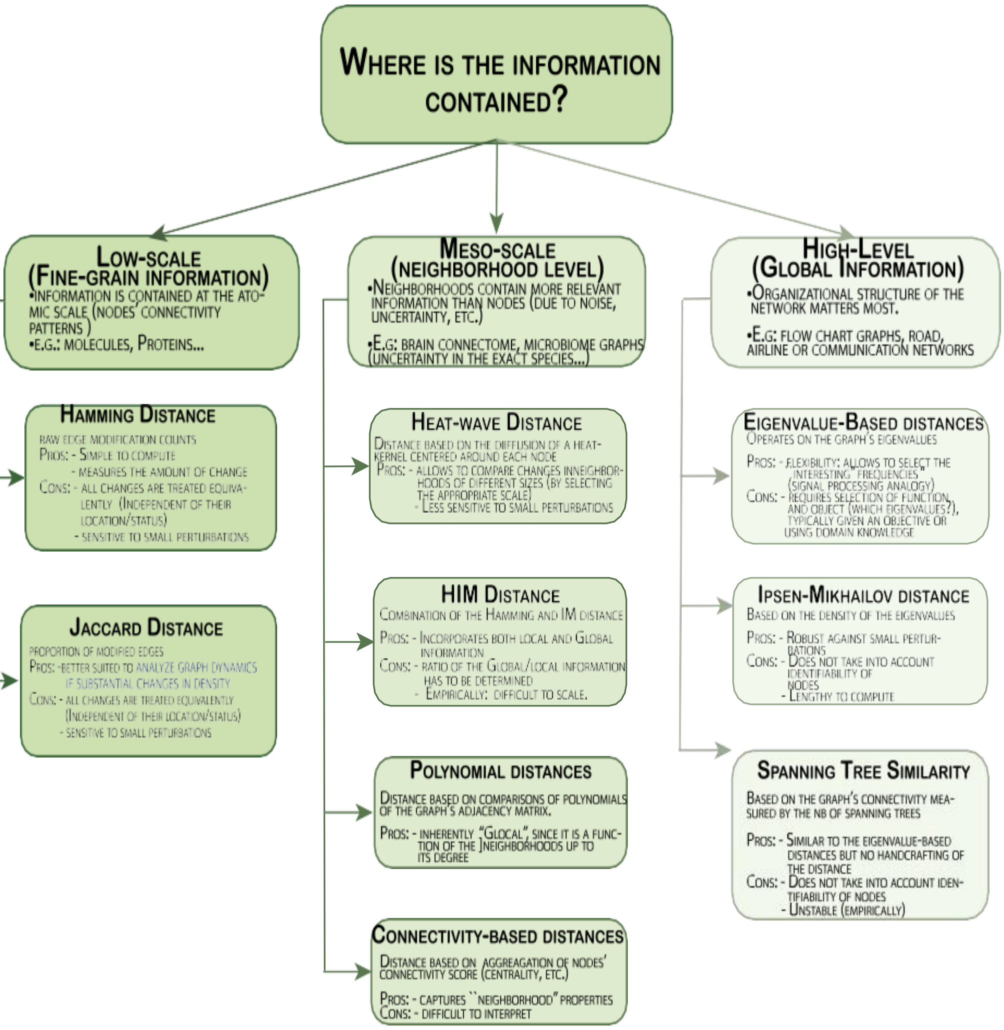
$$C = \begin{pmatrix} 0.4 & 0.1 & 0.001 \\ 0.1 & 0.2 & 0.01 \\ 0.001 & 0.01 & 0.5 \end{pmatrix}$$

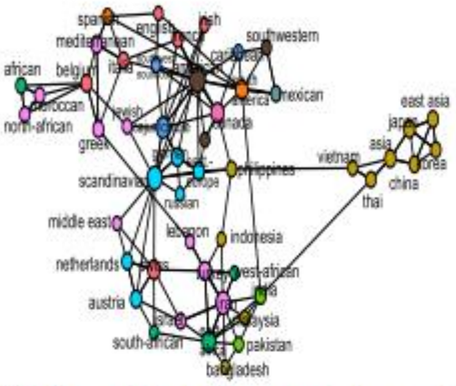
on the network analysis there topology is accessed by sets of network families in different global and local to

Stochastic Block Model topology - results

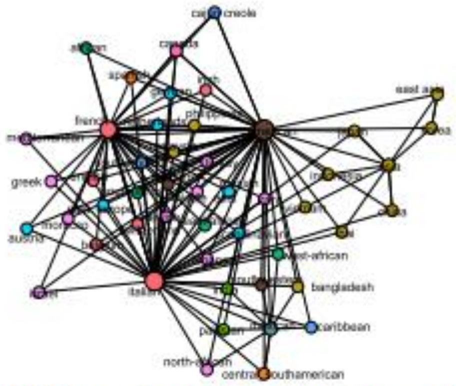
Row Top = Smooth dynamics is compared with 0.05% edges rewired at every time step.

Row Bottom = detection experiment of Change point

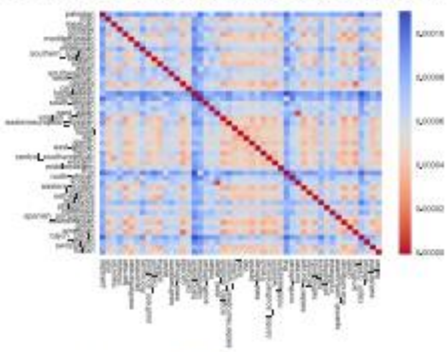




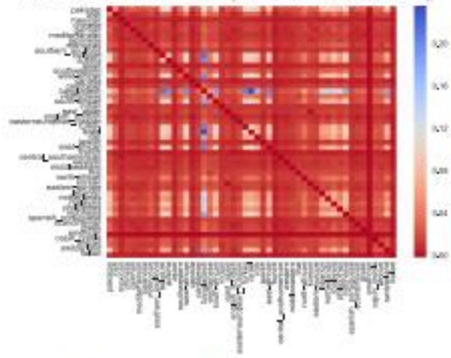
(a) 3-nearest-neighbor proximity graph between cuisines (Hamming distance)



(b) 3-nearest-neighbor proximity graph between cuisines (Jaccard distance)



(c) Pairwise distances between cuisines (Hamming distance)



(d) Pairwise distances between cuisines (Jaccard distance)

Conclusion

we have come acrossed so many graphs. For different node labels nodes never sets married. Easily graph distance and graph identities are very important. We have exampled the statistics by the use of metrics on graph objects.

In many sources distances are very essential to assess the variability in a dataset latent clusters (or) gradients are detected by multidimensional scaling embeddings of graphs in Euclidean space.

For instance, we have not used subgraphs and motif counts the same think which was quantifying in the past already. Some times we have used graph kernels.

References

[1] Pellegrini Matteo, Haynor David, Johnson JM: Protein interaction networks. Expert Rev Proteomics. 2004, 1 (2):

[2] Vikis HG, Guan KL: Glutathione-S-transferase-fusion based assays for studying protein-protein interactions. Methods Mol Biol. 2004, 261: 175-186.

[3] Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Seraphin B: The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods. 2001, 24 (3): 218-229.

[4] Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y: A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc Natl Acad Sci USA. 2001, 98 (8): 4569-4574.

[5] Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM: Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature. 2002, 415 (6868): 141-147.

[6] Stoll D, Templin MF, Bachmann J, Joos TO: Protein microarrays: applications and future challenges. Curr Opin Drug Discov Devel. 2005, 8 (2): 239-252.

[7] Willats WG: Phage display: practicalities and prospects. Plant Mol Biol. 2002, 50 (6): 837-854.

[8] Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M: Global mapping of the yeast genetic interaction network. Science. 2004, 303 (5659): 808-813.

[9] Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP: Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature. 2006, 440 (7084): 637-643.

[10] Xenarios I, Rice DW, Salwinski L, Baron MK, Marcotte EM, Eisenberg D: DIP: the database of interacting proteins. Nucleic Acids Res. 2000, 28 (1): 289-291.

[11] Mewes HW, Frishman D, Mayer KF, Munsterkotter M, Noubibou O, Pagel P, Rattei T, Oesterheld M, Ruepp A, Stumpflen V: MIPS: analysis and annotation of proteins from whole genomes in 2005. Nucleic Acids Res. 2006, D169-172. 34 Database.