# Study on the Clustering Coefficients in Metabolic Network Using a Hierarchical Framework

**Jose C. Nacher** nacher@kuicr.kyoto-u.ac.jp
**Nobuhisa Ueda** ueda@kuicr.kyoto-u.ac.jp
**Takuji Yamada** takuji@kuicr.kyoto-u.ac.jp
**Minoru Kanehisa** kanehisa@kuicr.kyoto-u.ac.jp
**Tatsuya Akutsu** takutsu@kuicr.kyoto-u.ac.jp

Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto 611-001, Japan

**Keywords:** scale-free network, self-organized systems, metabolic network, hierarchical organization

## 1 Introduction

Evidences for hierarchical organization in many real networks (biological and non-biological) have recently been reported [2, 3]. In particular, protein domain network and metabolic network have an inherently hierarchical topology. In order to reproduce the observed properties in real networks, *Ravasz et al.* proposed a hierarchical model with modular topology (the RSMOB model in what follows) [3]. These observed properties of networks with $N$ nodes are: scale-free of degree distribution $P(k) \sim k^{-\gamma}$, scaling of clustering coefficient $C(k)$ and high average clustering coefficient $C(N)$ that is independent of the network size (definitions of these three functions are found in Table 1).

In the metabolic network nodes are chemical compounds and edges are reactions. This network can also be understood as a bipartite graph. The network generated by only chemical compounds (reactions) is called compound (reaction) projection. A line graph transformation (i.e., each edge between two compounds becomes a node (reaction) of the transformed network) relates to both projections.

In this short abstract, we show that it is enough to apply the line graph transformation over the hierarchical network generated by the RSMOB model to evoke the clustering topology of the reaction network. A detailed description and mathematical derivation will be presented in a forthcoming work.

## 2 Method and Results

In the RSMOB model, the network is simultaneously scale-free and has high average clustering coefficient $C(N)$ that is independent of the network size. The network is made of densely connected 5-node modules that are assembled into larger 25-node modules by connecting the peripheral nodes to the central node of the original 5-node module (iteration n=1, $5^2 = 25$ nodes). In the next step four replicas are created to produce 125-node modules (iteration n=2, $5^3 = 125$ nodes). This process can be repeated indefinitely. Such modules (or building blocks) have been suggested to be a fundamental feature of biological networks [1]. The key signature of hierarchical modularity is the scaling of the clustering coefficient on the node degree $k$ as $\sim k^{-1}$.

Given an undirected graph $G$, defined by a set of nodes $V(G)$ and a set of edges $E(G)$, we associate another graph $L(G)$, called the line graph of $G$, in which $V(L(G)) = E(G)$, and where two vertices are adjacent iff they have a common endpoint in $G$ (i.e., $E(L(G)) = \{\{(u,v),(v,w)\}|(u,v) \in E(G),(v,w) \in E(G)\}$). This construction of graph $L(G)$ from the graph $G$ is called line graph transformation.

To evaluate the clustering coefficient of the reaction network ($C^T(k)$), we proceed in two steps. In the first step, we construct a hierarchical network with 4 initial nodes and up to 5 iterations (4096

nodes) by using RSMOB model. In the second step, we apply the line graph transformation on that network. The resulting network is called transformed network. In Fig. 1 (a), we show the clustering coefficient of this transformed network $C^T(k)$ with empty triangles. While for the initial hierarchical network (compound network) $C(k)$ scales as $k^{-1.1}$, for the transformed network (reaction network) $C^T(k)$ scales weakly as $k^{0.08}$. This theoretical prediction was compared with the experimental data from KEGG database finding a good agreement. In Fig. 1 (b), we plot $C(N)$ for networks with 3 and 4 initial nodes (filled circles and triangles). The values become asymptotically constant. The experimental value of the reaction network (star) is in good agreement with the asymptotic values of the transformed network $C^T(N)$ (empty triangles and circles).

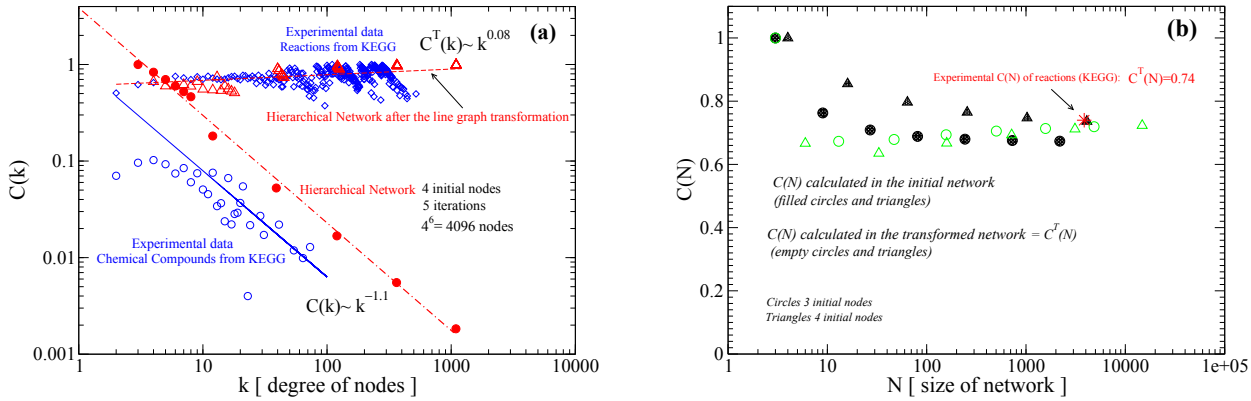


Figure 1: (a) Filled circles and dot-dashed line: $C(k)$ evaluated with the hierarchical network. Empty triangles and dashed line: $C^T(k)$. Diamonds: $C^T(k)$ of reactions data from KEGG. Empty circles and continuous line: $C(k)$ of compounds data from KEGG. (b) Filled circles and triangles: $C(N)$. Empty ones: $C^T(N)$. Star: Experimental $C^T(N)$ of reactions from KEGG (averaged over 163 organisms).

Table 1: Summary of functions and values before and after the line graph transformation is done. $n(k)$: number of nodes of degree $k$. $n_i$: number of edges connecting the $k_i$ nearest neighbours of node $i$ to each other. $C(k)$ is defined as the average clustering coefficient over nodes with the same degree $k$ of connectivity. $N$: Total number of nodes. The dagger symbol (†) means results obtained from our research.

| Function | Definition | Dependence *before* | Dependence *after* |
|---|---|---|---|
| $P(k)$ | $n(k)/N$ | $k^{-\gamma}$ | $k^{-\gamma+1}$ † |
| $C_i(k_i)$ | $2n_i/[k_i(k_i-1)]$ | | |
| $C(k)$ | $[\sum_{i:\,k_i=k} C_i(k_i)]/n(k)$ | $k^{-1.1}$ | $k^{0.08}$ † |
| $C(N)$ | $[\sum_i C_i(k)]/N$ | size-independent† | size-independent† |

## References


[1] Hartwell, L.H., Hopfield, J.J., Leibler, S. and Murray, A.W., From molecular to modular cell biology, *Nature*, 402:C47–C52, 1999.

[2] Ravasz, E. and Barabási, A.-L., Hierarchical organization in complex networks, *Phys. Rev. E*, 67:026112-1–7, 2003.

[3] Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N. and Barabási A.-L., Hierarchical organization of modularity in metabolic networks, *Science*, 297:1551–1555, 2002.