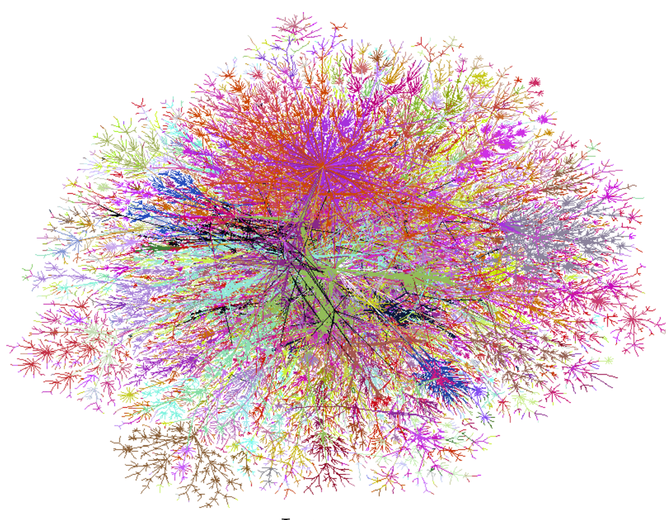
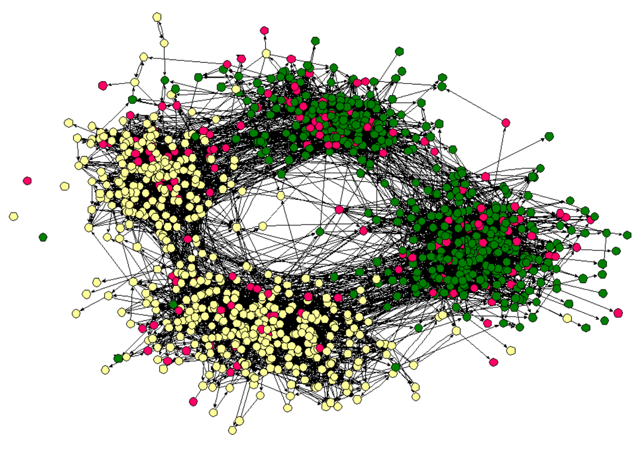
Distances in random Apollonian network structures

Olivier Bodini, Alexis Darrasse, and Michèle Soria

Motivations



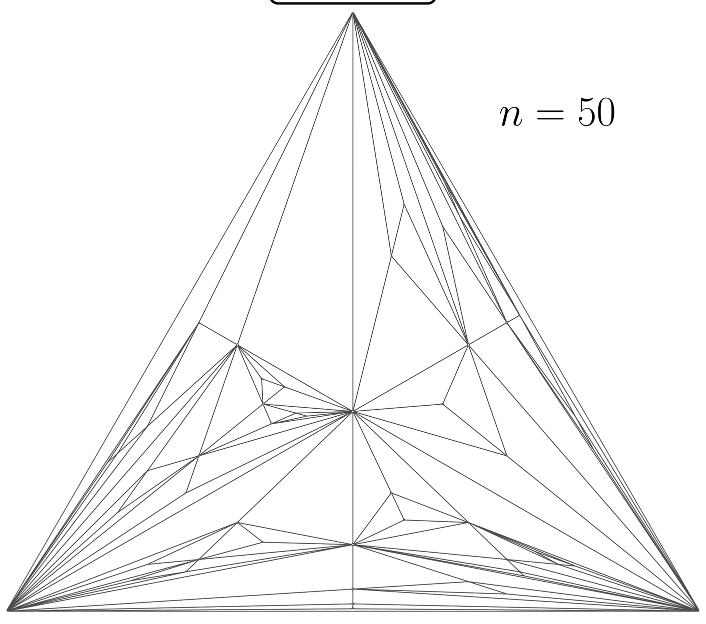
Internet
[Hal Burch and Bill Cheswick, Lumeta Corp.]



High school friendship
[James Moody, Race, school integration, and friendship segregation in America, American Journal of Sociology 107, 679-716 (2001)]

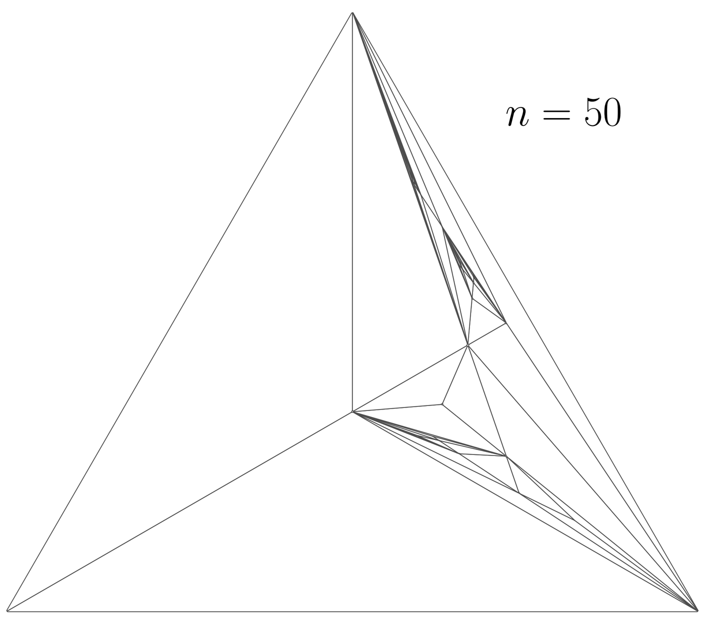
Real-world graphs have very different properties than those of random graphs (e.g. Erdös-Renyi): the degree distribution follows a power law, the mean distance is very small, etc.[2]

RAN [3]



Increasing ternary trees
Degree distribution: $P(k) \sim k^{-\gamma}$
Logarithmic distance

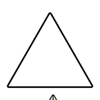
RANS



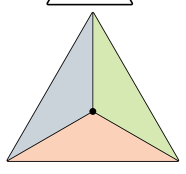
Ternary trees
Degree distribution: $P(k) \sim \left(\frac{8}{9}\right)^k k^{-3/2}$
Square-root distance

Definitions

A random Apollonian network structure (RANS) R is recursively defined as:



either an empty triangle,



or a triangle T split in three parts, by placing a vertex v inside T and connect it to the three vertices of the triangle; each sub-triangle being substituted by a RANS.

$\mathcal{O}(R) = \{O_1(R), O_2(R), O_3(R)\}$: the three vertices of the outermost triangle of RANS R.
$d(v, w)$: length of shortest path joining v to w.

References

[1] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*, web edition, 809+xii pages (available from the authors' web sites). To be published in 2008 by Cambridge University Press.

[2] M.E.J. Newman, A.L. Barabási and D.J. Watts. *The structure and dynamics of networks*. Princeton University Press, 2006

[3] T. Zhou, G. Yan, and B.-H. Wang. Maximal planar networks with large clustering coefficient and power-law degree distribution journal. *Physical Review E*, 71(4):46141, 2005.

Theorem 1

Given R a RANS of order n and v a random internal vertex of R, the distance from v to $O_1(R)$ has a Rayleigh limit distribution:

$$\Pr(d(v, O_1(R)) = x\sqrt{n}) = c\frac{x}{\sqrt{n}}e^{-c^2\frac{x^2}{4}}$$

and a mean value of $\frac{\sqrt{3\pi}}{11}\sqrt{n} + \frac{277}{363} + O(\frac{1}{\sqrt{n}})$.

Proposition

Multivariate generating function:

$$T_d(z, u_1, \dots, u_d) \equiv \sum r_{n,k_1,\dots,k_d} u_1^{k_1} u_2^{k_2} \dots u_d^{k_d} z^n, \quad r_{n,k_1,\dots,k_d}$$
$$= \#\{R \in \mathcal{R}_n \mid k_j \text{ vert. dist. } j \text{ from } O_1\}.$$

Recurrence relation:

$$T_d(z, u_1, \dots, u_d) = 1 + zu_1 T_d^2(z, u_1, \dots, u_d)\left(1 + zu_2\frac{1}{(1 - zu_2 T_{d-1}^2(z, u_2, \dots, u_d))^3}\right)$$

$$\text{and } T_1(z, u_1) = 1 + zu_1 T_1^2(z, u_1) T_0(z) \quad \text{with} \quad T_0(z) = T(z).$$

Lemma

Generating function for the number of vertices at distance i from O_1:
$D_i(z) = \frac{\partial}{\partial u_i} T_i(z, u_1, \dots, u_i)\Big|_{u_j=1,\forall j} = \sum_n k_i\, r_{n,k_i} z^n$.
D_i express as a function of z and $T(z)$:

$$D_{i+1}(z) = H^i(z) \times \frac{(1 + 2z^2T^4(z))}{6zT(z)(1 - 2zT^2(z))}, \quad \text{for } i \geq 2$$

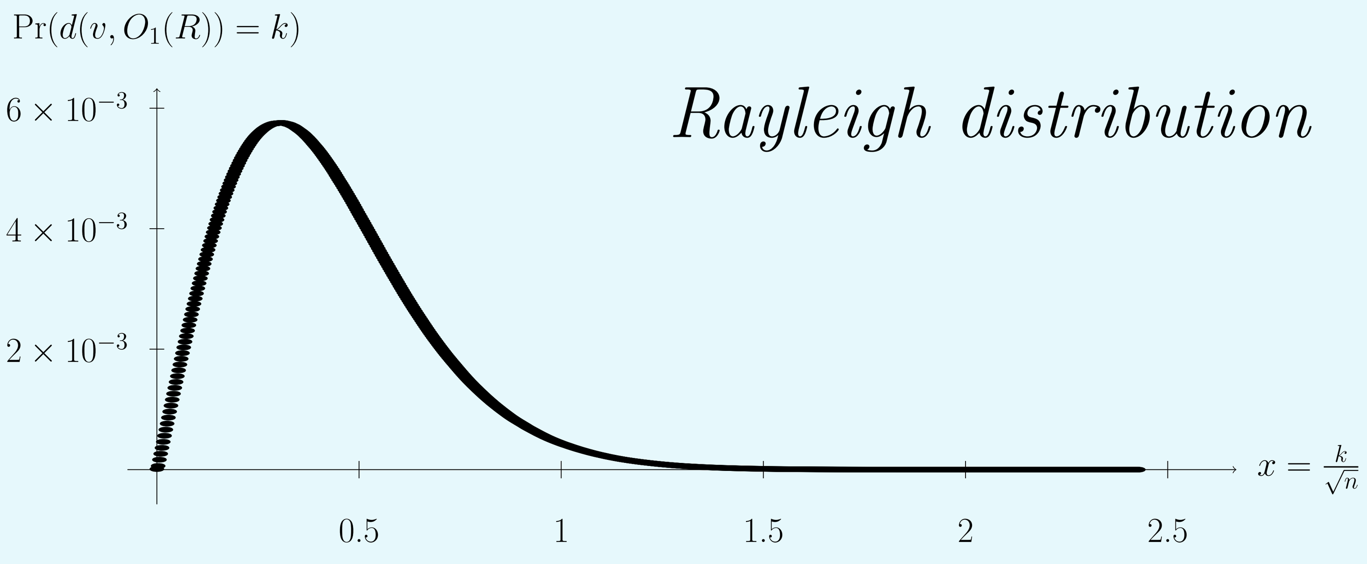
where $H(z) = 1 - \frac{11}{\sqrt{3}}\sqrt{1 - z/\rho} + \frac{2}{3}(1 - z/\rho) + (1 - z/\rho)^{3/2} + O((1 - z/\rho)^2)$, $\rho = 4/27$.

Sketch of proof

The full singular expansion of $D_i(z)$ can be derived from its expression in terms of H and D_2. Thus the proportion of vertices at distance i from O_1, that is $\frac{1}{nT_n}[z^n]D_i(z)$ can be evaluated:

$$\Pr(d(v, O_1(R)) = i) = \frac{1}{nT_n}[z^n]D_i(z) = \frac{1}{nT_n}[z^n]H^{i-2}(z)D_2(z).$$

The result follows from theorem IX.16 (Semi-large powers) of [1]: the singular exponent 1/2 for $H(z)$ implies a Rayleigh distribution for $k = x\sqrt{n}$.



Theorem 2

Let R be a RANS of order n and v, w two random vertices of R, the distance from v to w has mean value

$$E_{v,w\in R}(d(v, w)) = \frac{\sqrt{3\pi}}{11}\sqrt{n} + \frac{376}{363} + \frac{17\sqrt{3\pi}}{72}\frac{1}{\sqrt{n}} + \frac{25858246}{1185921}\frac{1}{n} + O(n^{-\frac{3}{2}}).$$

Definitions

Distances to one or two or three outermost vertices:

$$\Delta_{①}(R) = \sum_{x\in R} d(x, O_1(R)), \quad \Delta_{②}(R) = \sum_{x\in R} d(x, \{O_1(R), O_2(R)\}), \quad \Delta_{③}(R) = \sum_{x\in R} d(x, \mathcal{O}(\mathcal{R})).$$

Proposition

Multivariate generating function:

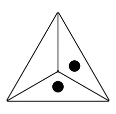
$$\Delta(z, d_①, d_②, d_③) \equiv \sum_{R\in\mathcal{R}} d_①^{\Delta_1(R)} d_②^{\Delta_2(R)} d_③^{\Delta_3(R)} z^{|R|} = \sum_{n,i,j,k=0}^{\infty} \alpha_{n,i,j,k}\, d_①^i d_②^j d_③^k z^n,$$

with $\alpha_{n,i,j,k} = \#$ RANS of order $n \mid \Delta_① = i, \Delta_② = j, \Delta_③ = k$.

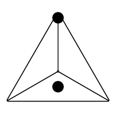
Recursive equation:

$$\Delta(z, d_1, d_2, d_3) = 1 + zd_①d_2d_③ \times \Delta(zd_①, d_2, d_3, d_1)$$
$$\times \Delta(z, d_①, d_2d_③, 1)$$
$$\times \Delta(z, d_1d_②, d_③, 1).$$

Families of pairs of vertices



interdistances — size n^2 — mean dist. $\sqrt{n}$

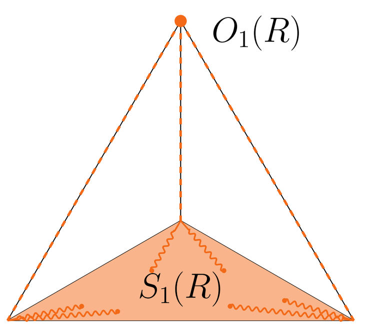


intradistances — size $n^{3/2}$ — mean dist. $\sqrt{n}$

Intradistances

Top level:

$$\delta(z) = \sum(3 + \Delta_3(S_1(R)) + |S_1(R)|$$
$$+ \Delta_3(S_2(R)) + |S_2(R)|$$
$$+ \Delta_3(S_3(R)) + |S_3(R)|)\; z^{|R|}$$
$$= 3T(z) + 3zT^2(z)\Delta_3(z) + 3z^2T^2(z)T'(z)$$



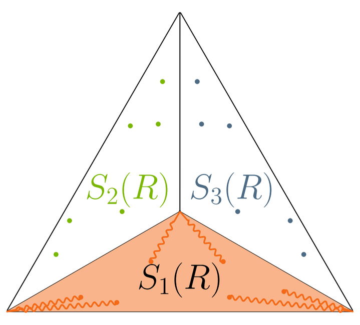
Recursively:

$$\text{Intra}(z) = \delta(z)\frac{T'(z)}{T^3(z)} \sim 3z\Delta_3(z)T'(z)/T(z) \Rightarrow \frac{[z^n]\text{Intra}(z)}{[z^n]T(z)} \sim \frac{1}{44}n^2$$

Interdistances

Top level:

$$\gamma^-(z) = 3\sum_{R\in\mathcal{R}} \Delta_3(S_1(R)) \times (|S_2(R)| + |S_3(R)|)\, z^{|T|}$$
$$= 6z^2T(z)T'(z)\Delta_3(z)$$



Recursively:

$$\text{Inter}^-(z) = \gamma^-(z)\frac{T'(z)}{T^3(z)} = 6\Delta_2(z)z^2T'^2(z)/T^2(z) \Rightarrow \frac{[z^n]\text{Inter}^-(z)}{[z^n]T(z)} \sim \frac{\sqrt{3\pi}}{22}n^2\sqrt{n}$$