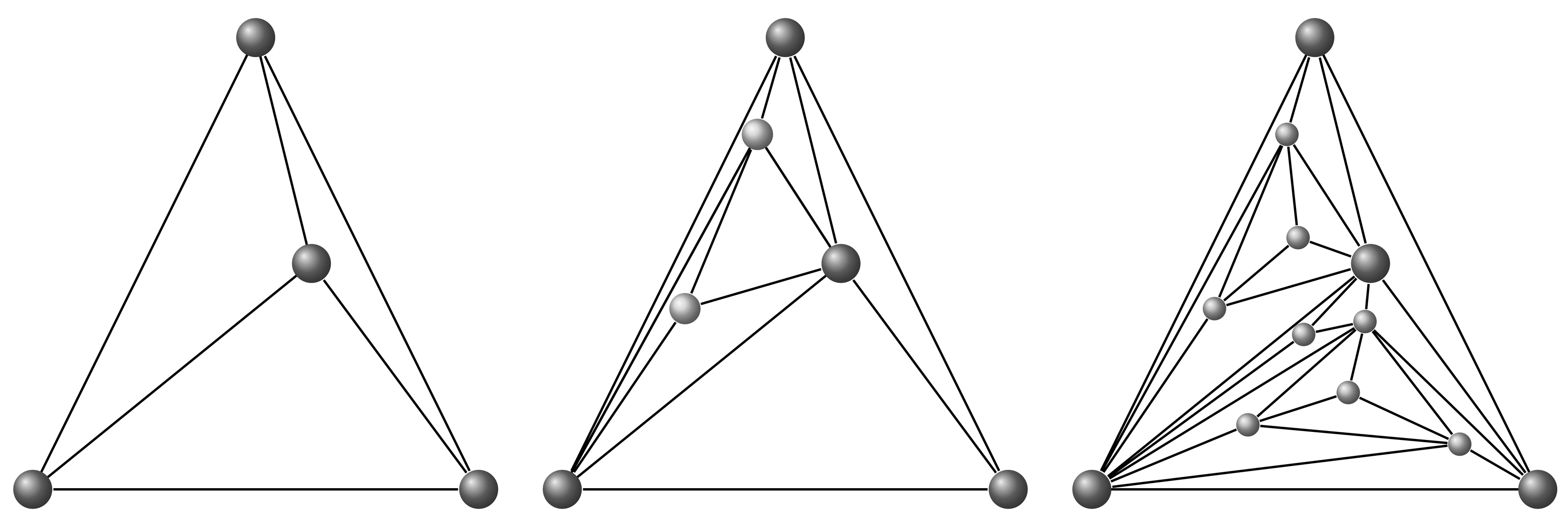
Properties of Random Apollonian Networks

István Kolossváry, Lajos Vágó, joint with Júlia Komjáthy
Budapest University of Technology, Stochastics Department

Definition of Random Apollonian Networks

- Initially in d dimensions start with a K_{d+2}. Active cliques are the $d+1$-cliques.
- At step $n \geq 1$, pick an active clique C uniformly at random, which becomes inactive.
- Insert new node. All possible $d+1$-cliques with vertices of C become new active cliques.
- Repeat. Result after n steps is a $\mathrm{RAN}_d(n)$. Figure shows a $\mathrm{RAN}_2(n)$ for $n = 0, 2, 8$.



Real-life propeties

- Power law degree distribution: this is a kind of preferential attachment model. The larger the degree, the more faces are adjacent to it, so its degree grows with greater probability.
- Large clustering coefficient: in every step we are creating cliques.
- Small world: a hierarchical structure unfolds, uncovering a tree-like structure, whose branches are pretty evenly distributed in depth.
- We give precise formulations and proofs for these properties.

Degree distribution, clustering coefficient

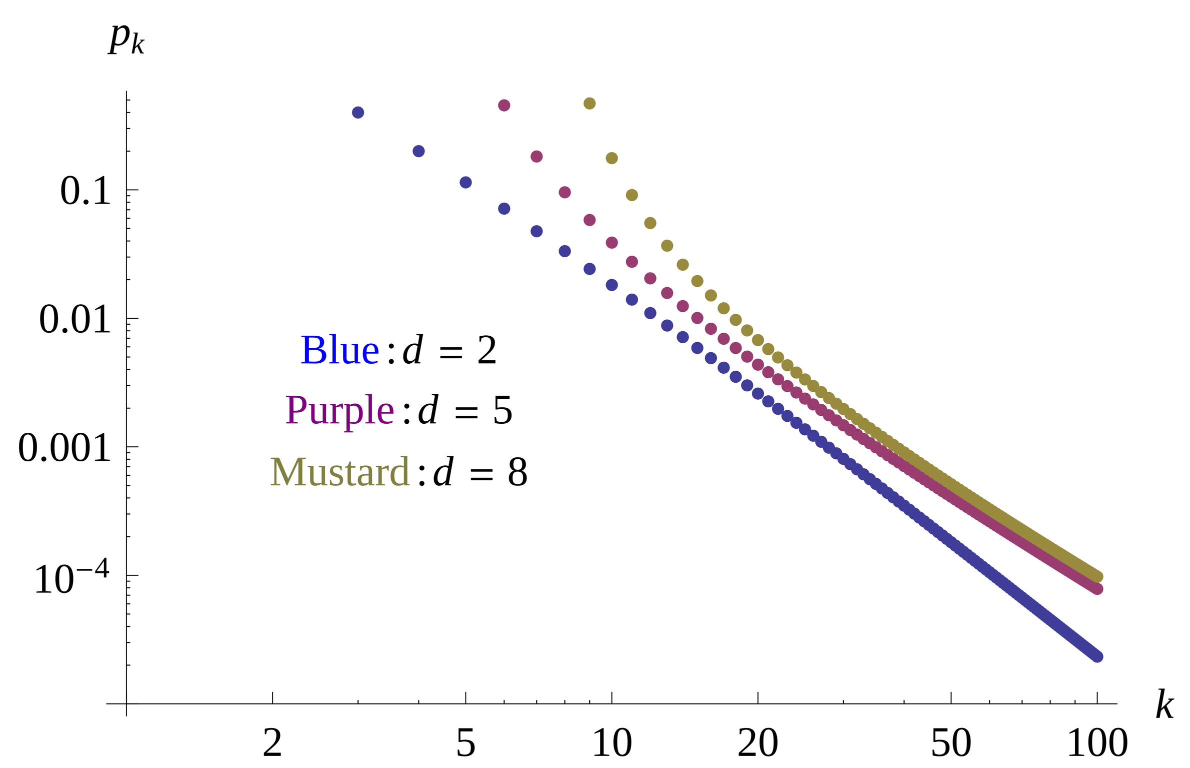
- $N_k(n) := \#\{$vertices with degree k at time $n\}$
- $p_k(n) := \{$proportion of vertices with degree k at time $n\} = N_k(n)/|V(n)|$, where $|V(n)|$ stands for the total number of vertices.

Theorem for degree distribution

There exists a probability distribution $\{p_k\}_{k=d+1}^{\infty}$ and a constant c for which

$$\mathbf{P}\left(\max_k |p_k(n) - p_k| \geq c\sqrt{\frac{\log n}{n}}\right) = o(1).$$

Further, p_k follows a power law with exponent $(2d-1)/(d-1) \in (2,3]$.



The clustering coefficient of a vertex with degree k is *deterministic* and equals

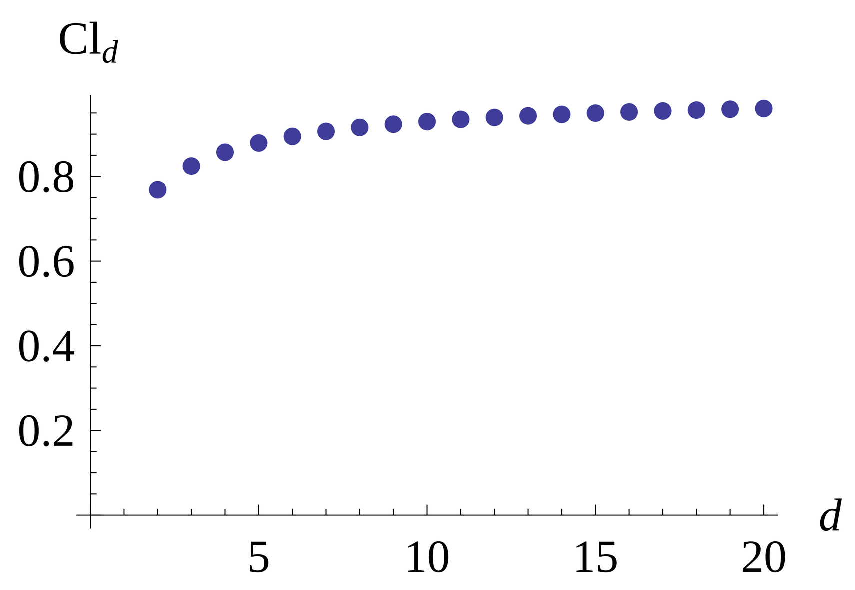
$$\frac{d(2k-d-1)}{k(k-1)} \sim \frac{2d}{k}.$$

Hence the clustering coefficient is

Corollary: Clustering coefficient of RANs

The average clustering coefficient of $\mathrm{RAN}_d(n)$ converges to a strictly positive constant as $n \to \infty$, given by

$$Cl_d = \Sigma_{k=d+1}^{\infty} \frac{d(2k-d-1)}{k(k-1)} p_k.$$



References

I.K., J.K. and L.V. Degrees and distances in random and evolving Apollonian networks. arXiv:1310.3864 [math.PR], 2013.

Z. Zhang, F. Comellas and L. Rong. High-dimensional random Apollonian networks. *Physica A: Statistical Mechanics and its Applications*, 2006.

A. Frieze and C. E. Tsourakakis. On certain properties of random Apollonian networks. In Proceedings of the 9th int. conf. on Algorithms and Models for the Web Graph, 2012.

M. Albenque and J-F. Marckert. Some families of increasing planar maps. *Electron. J. Probab.*, 2008.

B. Bollobás, O. Riordan, J. Spencer and G. Tusnády. The degree sequence of a scalefree random graph process. *Random Structures Algorithms*, 2001.

Shortest paths

$\mathrm{Hop}_d(n,u,v) = \#\{$edges on shortest path between u and $v\}$,
$\mathrm{Flood}_d(n,u) = \max_v \mathrm{Hop}_d(n,u,v)$ and $\mathrm{Diam}_d(n) = \max_{u,v} \mathrm{Hop}_d(n,u,v)$.
Coupon collector random variable $Y_d := \Sigma_{i=1}^{d+1} \mathrm{Geo}(\frac{i}{d+1})$, $\mu_d := \mathbf{E}[Y_d]$, $\sigma_d^2 := \mathbf{Var}[Y_d]$.

Theorem for hopcount, flooding time and diameter

The hopcount between the vertices of two uniformly chosen active cliques satisfies

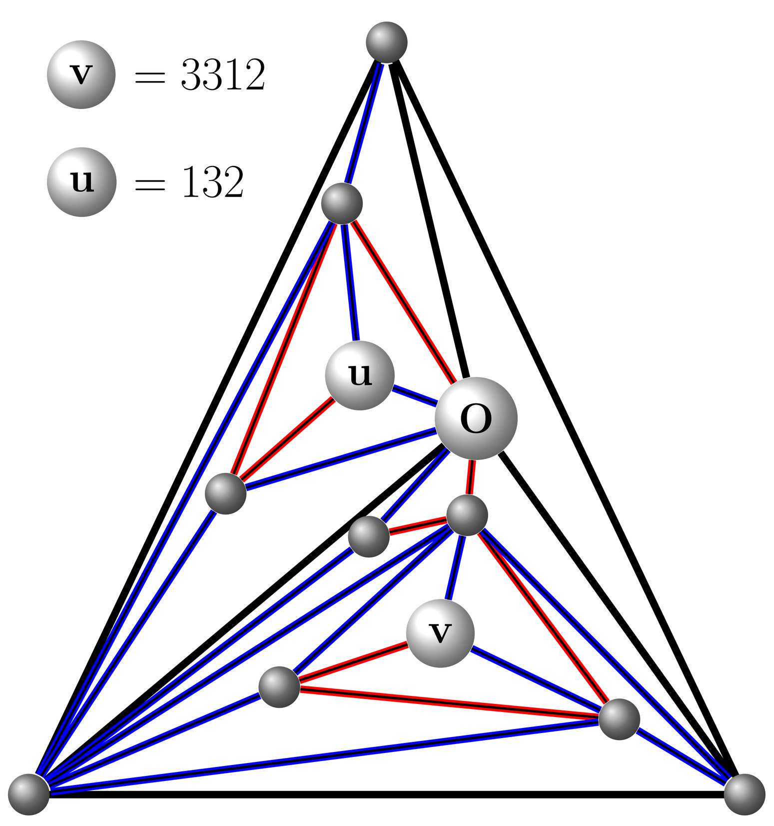
$$\frac{\mathrm{Hop}_d(n) - \frac{2}{\mu_d}\frac{d+1}{d}\log n}{\sqrt{2\frac{\sigma_d^2+\mu_d}{\mu_d^3}\frac{d+1}{d}\log n}} \xrightarrow{d} Z, \text{ where } Z \sim \mathcal{N}(0,1).$$

Let $f_d(c) := c - \frac{d+1}{d} - c\log(\frac{d}{d+1}c)$ and define $\tilde{c}_d = \{c_d > \frac{d+1}{d} : \ f_d(c_d) = -1\}$.
Then

$$\frac{\mathrm{Flood}_d(n)}{\log n} \xrightarrow{\mathbf{P}} \frac{1}{\mu_d}\left(\frac{d+1}{d} + \tilde{\alpha}\tilde{\beta}\tilde{c}_d\right) \quad \text{and} \quad \frac{\mathrm{Diam}_d(n)}{\log n} \xrightarrow{\mathbf{P}} 2\tilde{\alpha}\tilde{\beta}\frac{\tilde{c}_d}{\mu_d},$$

where $(\tilde{\alpha}, \tilde{\beta})$ maximize $\alpha\beta$ under the constraint $1 + f_d(\alpha\tilde{c}_d) - \alpha\beta\frac{\tilde{c}_d}{\mu_d} I_d\left(\frac{\mu_d}{\beta}\right) = 0$.

Tree-like structure of RANs

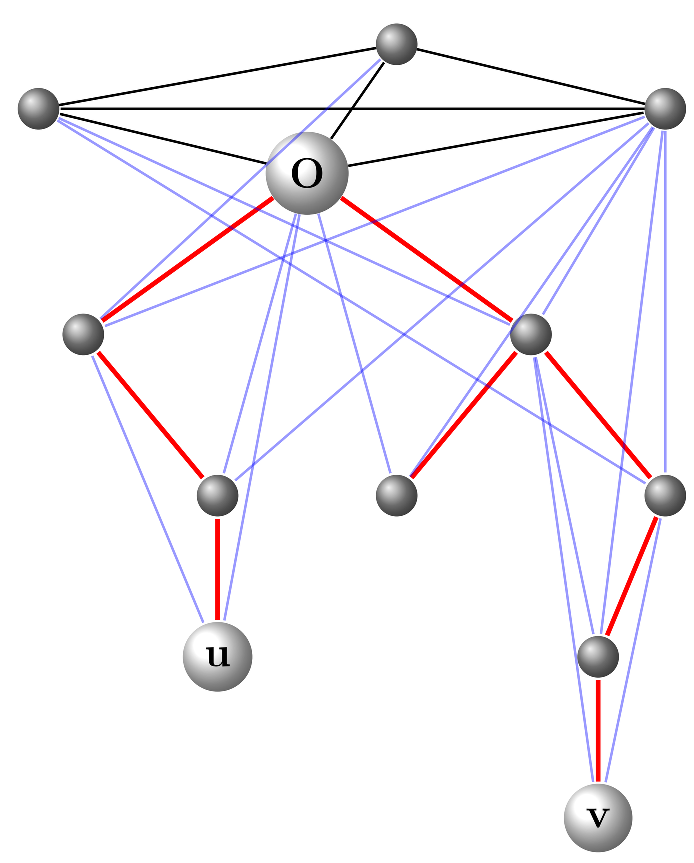


Coding of vertices

- Alphabet $\Sigma_d = \{1, 2, \ldots, d+1\}$
- $d+1$ Neighbors of new vertex $\mathbf{v}$: drop subcode from end until last occurrence of $i \in \Sigma_d$ in $\mathbf{v}$.

— Initial graph $\sim$ root of tree
— Forward edges $\sim$ branches of tree
— Shortcut edges along a branch
Go down the hierarchy along forward edges and climb back up faster with shortcut edges.

- Fastest way to reach root by longest hops: shortest suffix where each symbol appears at least once. Divide codes into such blocks. Ex. $113213323122 \to 1|132|1332|3122$.
- Deepest common ancestor $\mathbf{u} \wedge \mathbf{v}$. Shortest path intersects path between $\mathbf{O}$ and $\mathbf{u} \wedge \mathbf{v}$.



Main questions

- Length of the deepest common ancestor?
- Length of a typical / the longest code?
- Speed of ascent with shortcut edges?
- Large deviation for shortcut edges on almost longest code?

Diameter - sketch of proof

- Active clique chosen uniformly at random $\sim$ continuous time branching process with i.i.d. exponential lifetimes and deterministic offspring distribution $d+1$.
- Time when $\mathbf{u} \wedge \mathbf{v}$ splits has a limiting distribution. $\Rightarrow$ $\mathbf{u} \wedge \mathbf{v}$ is "close" to root
- $i \in \Sigma_d$ uniformly distributed among digits $\Rightarrow$ Longest hop $\sim$ coupon collector Y_d.
- G_m = generation of m-th chosen vertex = sum of indicators $\Rightarrow$ satisfies a CLT.
  $H_k = \#\{$full coupon collector blocks in a code of length $k\}$. Also satisfies a CLT.
  In the example 113213323122, $G_m = 12$ and $H_k = 3$.
- Large deviations for G_m and H_k:

$$\lim_{m\to\infty} \frac{\log\left(\mathbf{P}(G_m > c\log m)\right)}{\log m} = f_d(c), \quad \lim_{k\to\infty} \frac{1}{k}\log\left(\mathbf{P}\left(H_k > \frac{\beta}{\mu_d}k\right)\right) = -\frac{\beta}{\mu_d} I_d\left(\frac{\mu_d}{\beta}\right),$$

  where $1 \leq \beta \leq \mu_d/(d+1)$ and I_d is the rate function of Y_d. For the deepest branch

$$\frac{\max_{i\leq m} G_i}{\log m} \xrightarrow{\mathbf{P}} \tilde{c}_d, \quad \text{where } \tilde{c}_d \text{ as in the Theorem.}$$

- Entropy vs. energy argument. Deepest branch may not contain the most copies of Y_d $\Rightarrow$ count vertices u s.t. their depth $G_u > \alpha\tilde{c}_d \log m$ $(\alpha < 1)$, but have more than expected number of coupon collector blocks, i.e. $H_{G_u} > \frac{\beta}{\mu_d}\alpha\tilde{c}_d \log m$.
  Pick (α, β) so that the expected number of such vertices is constant, this yields the constraint $1 + f_d(\alpha\tilde{c}_d) - \alpha\tilde{c}_d\frac{\beta}{\mu_d} I_d\left(\frac{\mu_d}{\beta}\right) = 0$, under which we maximize $\frac{\beta}{\mu_d}\alpha\tilde{c}_d$.

Prepared for the Stochastic Activity Month taking place at EURANDOM.

This project was partially supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001**project. This project has been supported by the European Union, co-financed by the European Social Fund.

January 2014.