

# Fast Counting of Triangles in Large Real Networks

1. What is the problem that the paper wants to solve? Why is it difficult (related works)?
  - No. of triangles may help identifying anomalies in a graph, such as spam emails etc.
  - Given a graph authors want to find the numbers of triangles in the graph without actually counting them as counting is slow and should be avoided
2. What is the solution? What is the main idea?
  - Determine number of triangles only using the graph's eigenvalues
  - Number of triangles can be approximated using  $\Delta(G) = \frac{1}{6} \sum_{i=1}^n \lambda_i^3$
  - Number of triangles that a node  $i$  participates in is simply approximated by  $\Delta_i = \frac{1}{2} \sum_j \lambda_j^3 u_{i,j}^2$
  - Accurate because real-world graph's eigenvalues follow a power-law, alternating signs, top- $k$  largest eigenvalues can be used for approximation
3. What is the result?
  - Eigenvalues can be computed much faster, algo. easily parallelizable (map-reduce)
  - Big speedup in computation of number of triangles
  - Only 6.2 largest eigenvalues needed for accuracy  $\geq 95\%$  (mean, std. 3.2)
  - Authors made four main observations for the number of triangles in large graphs:
    1. count of triangles a node participates in follows a power-law
    2.  $\Delta_{avg}^{d_m}$ , average number of triangles of nodes of degree  $i$ , follows a power-law
    3. slope of  $\Delta_{avg}^{d_m}$ -power-law is complementary to slope of degree distribution (if it follows a power-law)
    4. high degree nodes deviate from  $\Delta_{avg}^{d_m}$ -power-law as they have a lot of neighbors of degree 1
4. What is the main novelty that enabled the solution?
  - Main novelty is the theorems for the (global and local) number of triangles
  - Allows for the approximation by largest eigenvalues
5. What are the good aspects of the paper? Did you learn something from the paper?
  - Idea is conveyed in a very straightforward and clear manner, paper is easy to follow
  - Provided an algorithm that is parallelizable and therefore applicable to even very large graphs
6. What is the impact of the paper?
  - Authors provided an entirely new and much more efficient way to count triangles by reducing the problem to a linear algebra problem
7. Are there weaknesses/missing parts in the paper? How can you improve it?
  - It is not clear what kind of datasets the laws identified by the authors apply to beyond the datasets in the paper (more importantly, in which kind of datasets the laws do not apply)
8. How can you extend the paper?
  - Look at graphs of other sizes than the ones tested in the paper to make more general observations on the laws stated by the authors
9. How can you apply the technique to other data/problems?
  - The technique is applicable to any data with triangles for an approximation of the number of triangles