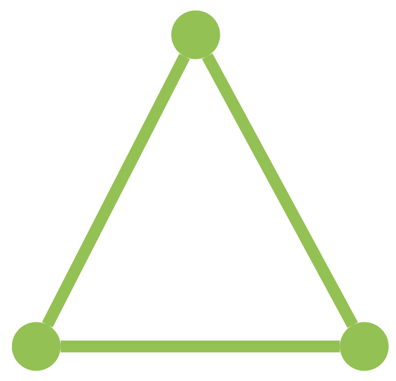
COUNTING TRIANGLES IN GRAPHS

Edina Marica

Supervisor: Ass.-Prof. Dr.techn. Gramoz Goranci

universität wien

INTRODUCTION



Triangles are:
- the simplest forms of **cliques** in graphs
- crucial for **community detection**
- essential for **pattern recognition**
- key for **social network analysis**

but they are **computationally expensive**.

GOAL

Compare different methods, **combinatorial** and **algebraic, exact and approximation** algorithms to analyze their performance and runtime.

DATASETS

Brain networks

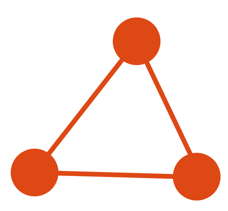
Social media networks

Co-purchase networks

Collaboration networks

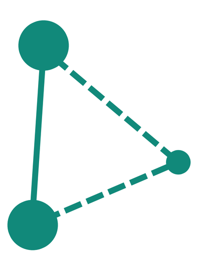
COMBINATORIAL ALGORITHMS

Algorithm 1 - naive method



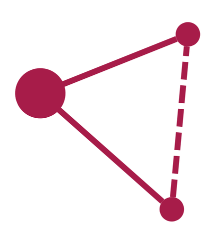
Take all node triplets, check if they are connected
Time complexity: $O(n^3)$

Algorithm 2 - edge iterator



Take two connected nodes, find common neighbor
Time complexity: $O(nm)$, smart way: $O(m^{1.5})$

Algorithm 3 - node iterator



Take a node, find neighbor pairs that are connected
Time complexity: $O(nm)$, smart way: $O(m^{1.5})$

EXPERIMENTAL RESULTS

Combinatorial algorithms						
Dataset	Vertices	Edges	Node it.	Fast Node it.	Edge it.	Fast Edge it.
Brain	213	16,089	0.28	0.05	0.27	0.06
Wiki	2,277	31,371	0.38	0.05	0.42	0.05
Relativity	5,242	14,484	0.04	0.02	0.04	0.02
Astrophysics	18,772	198,050	1.90	0.48	2.31	0.42
Email	36,692	183,831	2.94	0.41	3.13	0.41
Amazon	334,863	925,872	3.67	1.94	3.10	2.53
Twitch	168,114	6,797,557	-	101.34	-	36.70

ALGEBRAIC ALGORITHMS

Trace of the adjacency matrix

The number of triangles in an undirected graph is equal to $\frac{1}{6}tr(A^3)$.
Time complexity to calculate A^3: $O(n^3)$.

OBSERVATION: time complexity of a matrix-vector multiplication is $O(n^2)$, so we can calculate $A^3x = A(A(Ax))$ with $3n^2$ operations -> matrix-free method.

Approximating the trace

Algorithm 4 - Hutch

$H(A) = \frac{1}{m}\sum_{i=1}^{m} g_i^T A g_i$ -> m matrix-vector multiplications
- if $m = O(\frac{1}{\varepsilon^2})$ then $H(A)$ is an **ε-approximation** for $tr(A)$.

Algorithm 5 - Hutch ++

More sophisticated version of Hutch that requires only $m = O(\frac{1}{\varepsilon})$ matrix-vector multiplications.

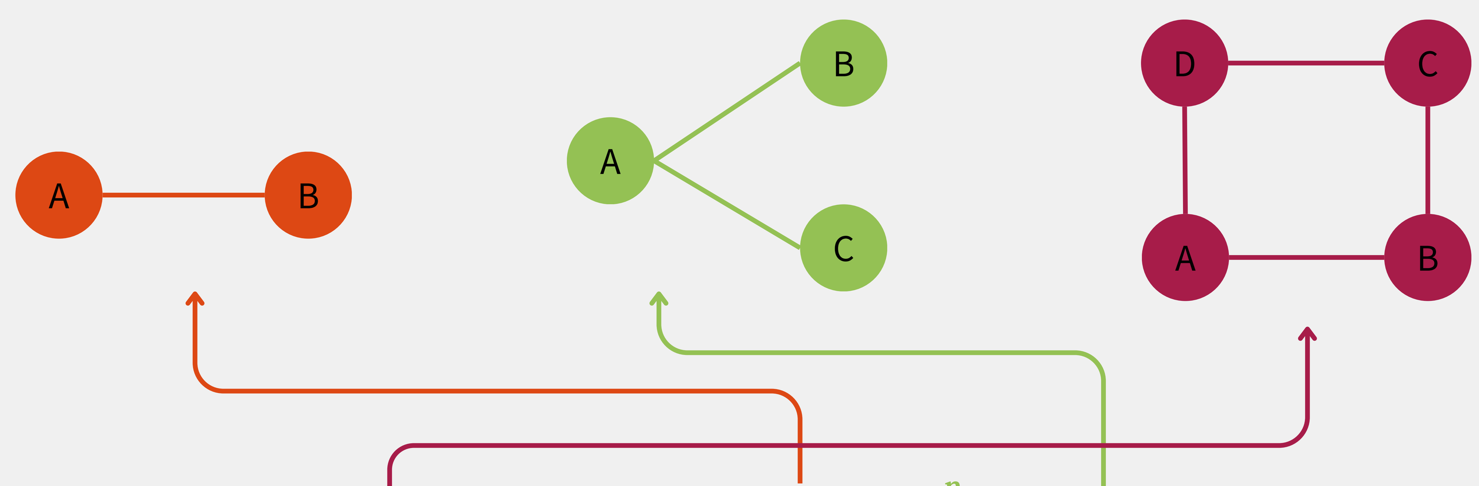
Algorithm 6 - Eigen Triangle

The trace can also be expressed with the eigenvalues of the adjacency matrix $tr(A) = \frac{1}{6}\sum_{i=1}^{n}\lambda_i^3$
- $tr(A)$ can be well approximated with the first eg. 30 eigenvalues.

EXPERIMENTAL RESULTS

Dataset	Triangle count	Fast Node	Fast Edge	Hutch++	
				time	relative error
Brain	622,414	0.05	0.06	0.004	0.0002
Wiki	343,066	0.05	0.05	0.009	0.002
Relativity	48,260	0.02	0.02	0.009	0.017
Astrophysics	1,351,441	0.48	0.42	0.06	0.069
Email	727,044	0.41	0.41	0.08	0.029
Amazon	667,129	1.94	2.53	1.49	0.094
Twitch	54,148,895	101.3	36.7	3.90	0.043

EXTENSION TO 4-CYCLES



$$N(G) = \frac{1}{8}\left(tr(A^4) - 2|E| - 2\sum_{i=1}^{n} d_i(d_i - 1)\right)$$

REFERENCES

- T. Schank, D. Wagner: Finding, Counting and Listing all Triangles in Large Graphs, An Experimental Study
- Chiba, Nishizeki: Arboricity and Subgraph Listing Algorithms
- M.F.Hutchinson: A Stochastic Estimator of the Trace of the Influence Matrix for Laplacian Smoothing Splines
- R. A. Meyer et al.: Hutch++: Optimal Stochastic Trace Estimation
- C. E. Tsourakakis: Fast Counting of Triangles in Large Real Networks: Algorithms and Laws