ESCAPE: Efficiently Counting All 5-Vertex Subgraphs

Authors: Ali Pinar, C. Seshadhri, Vaidyanathan Vishal
Presented by: Omar Obeya

Goal

- Count 5-vertex subgraphs
  - Exact
  - Scales

New Problem: 21 5-vertex patterns

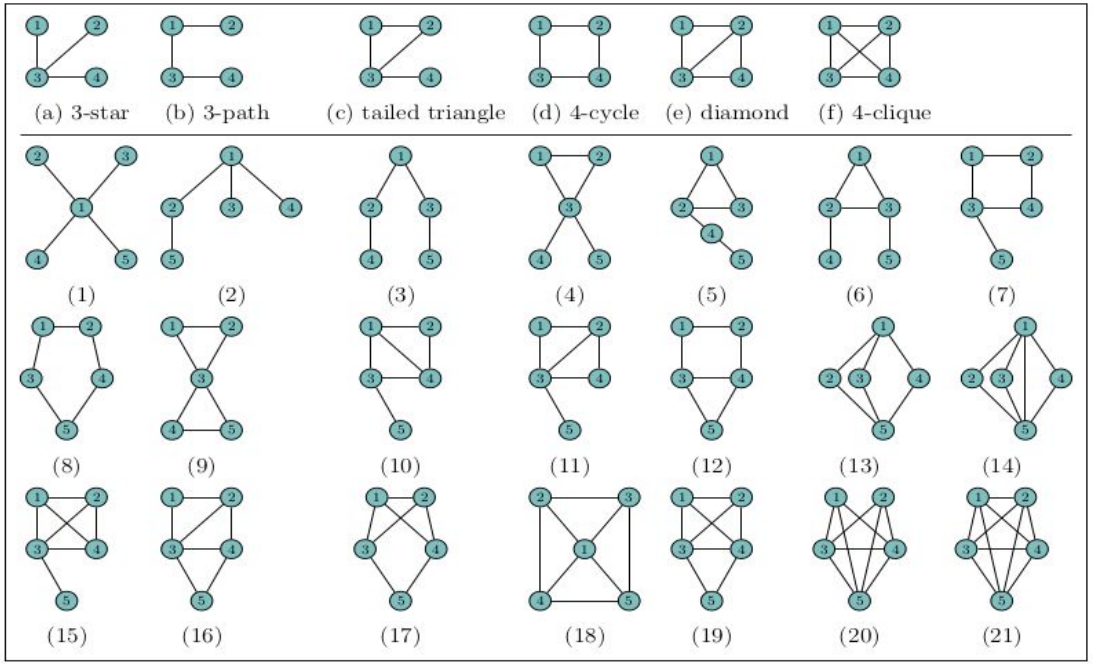
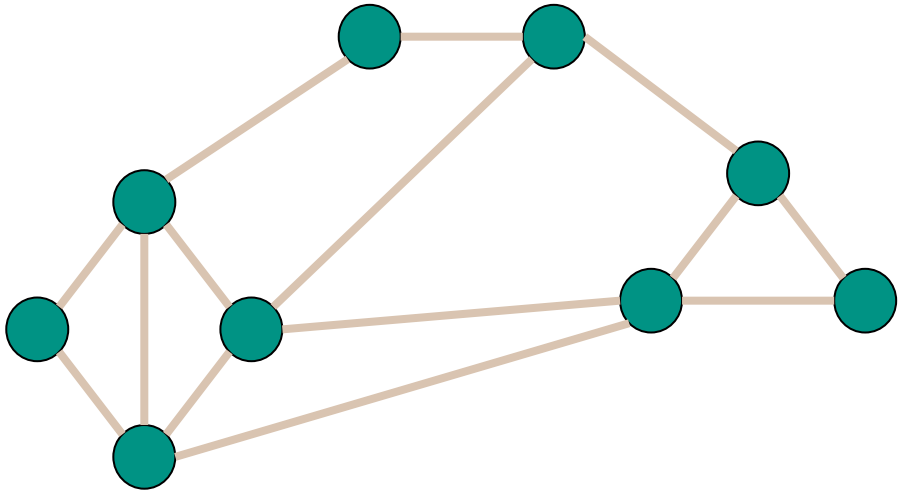
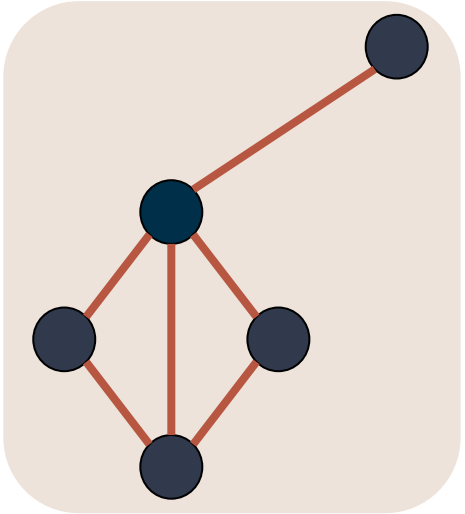


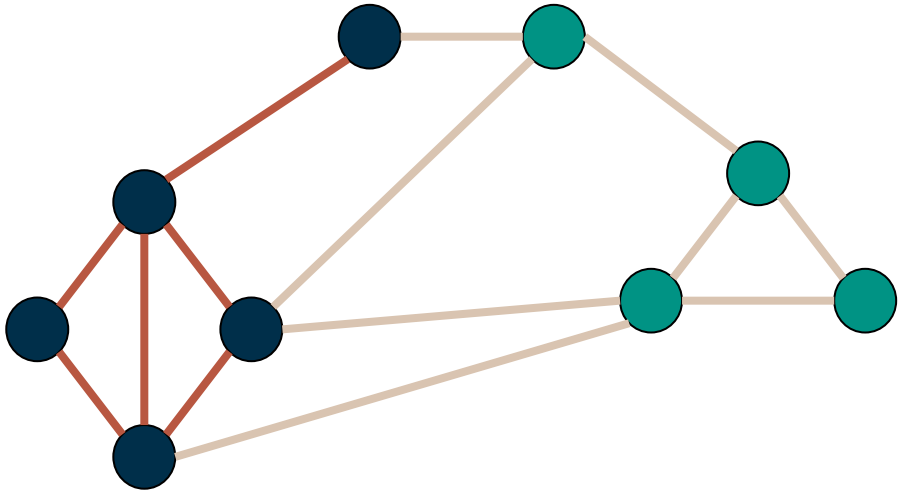
Figure 1: Connected 4 and 5-vertex patterns

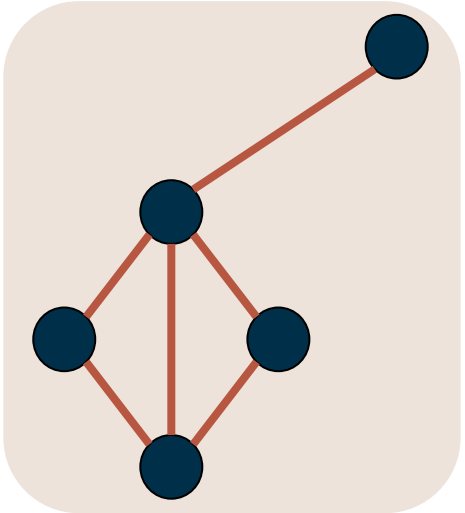
Counting Patterns



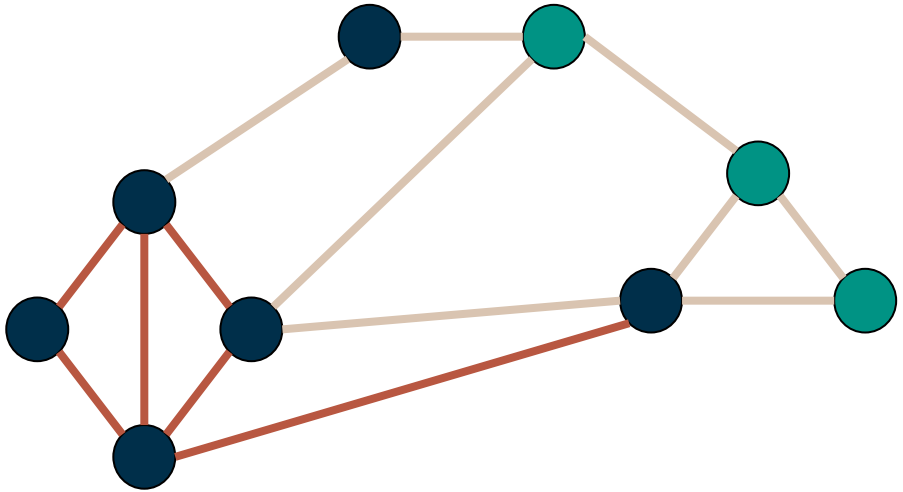


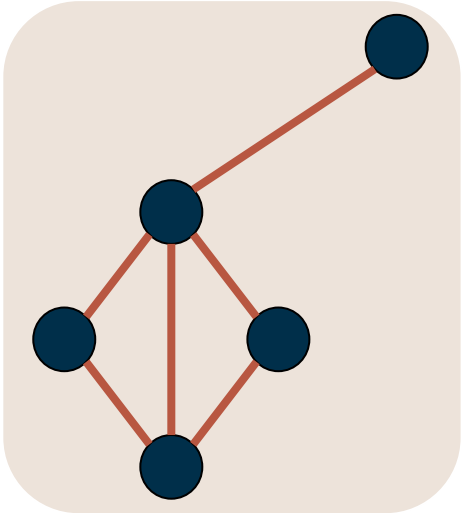
Counting Patterns





Counting Patterns

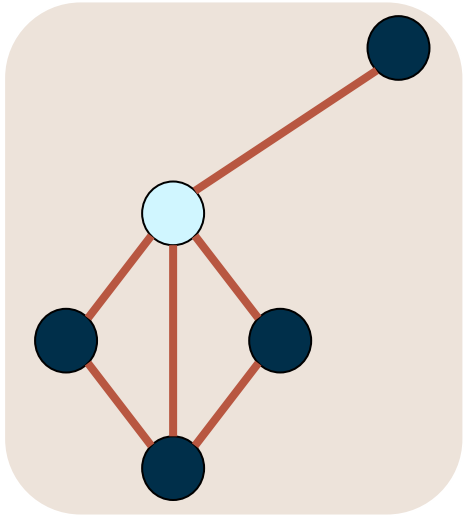


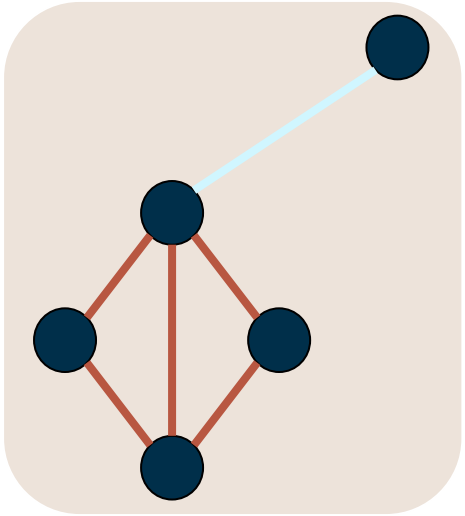


The Basics

Key Idea: Cutting

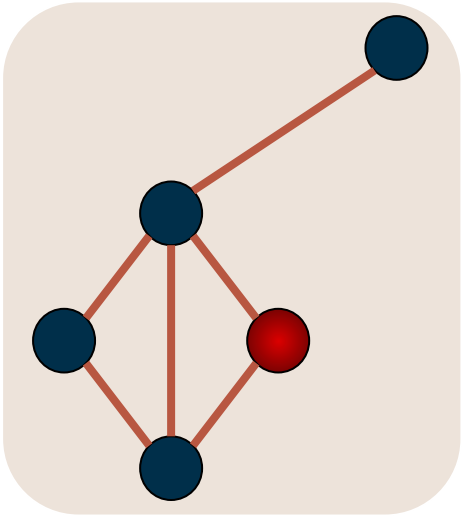
- A Cut



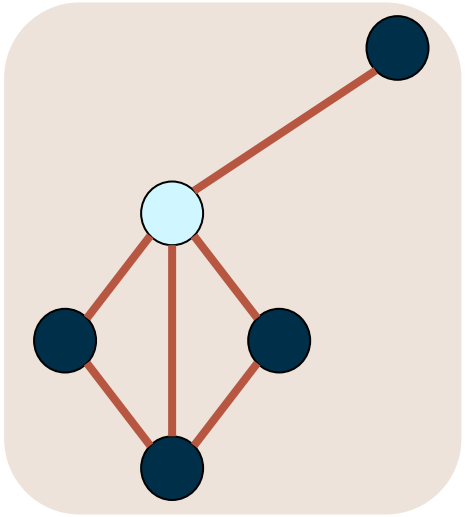


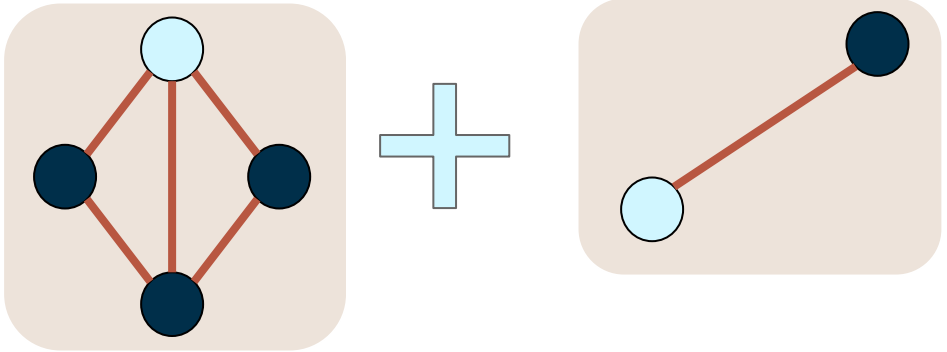
Key Idea: Cutting

- Not a Cut

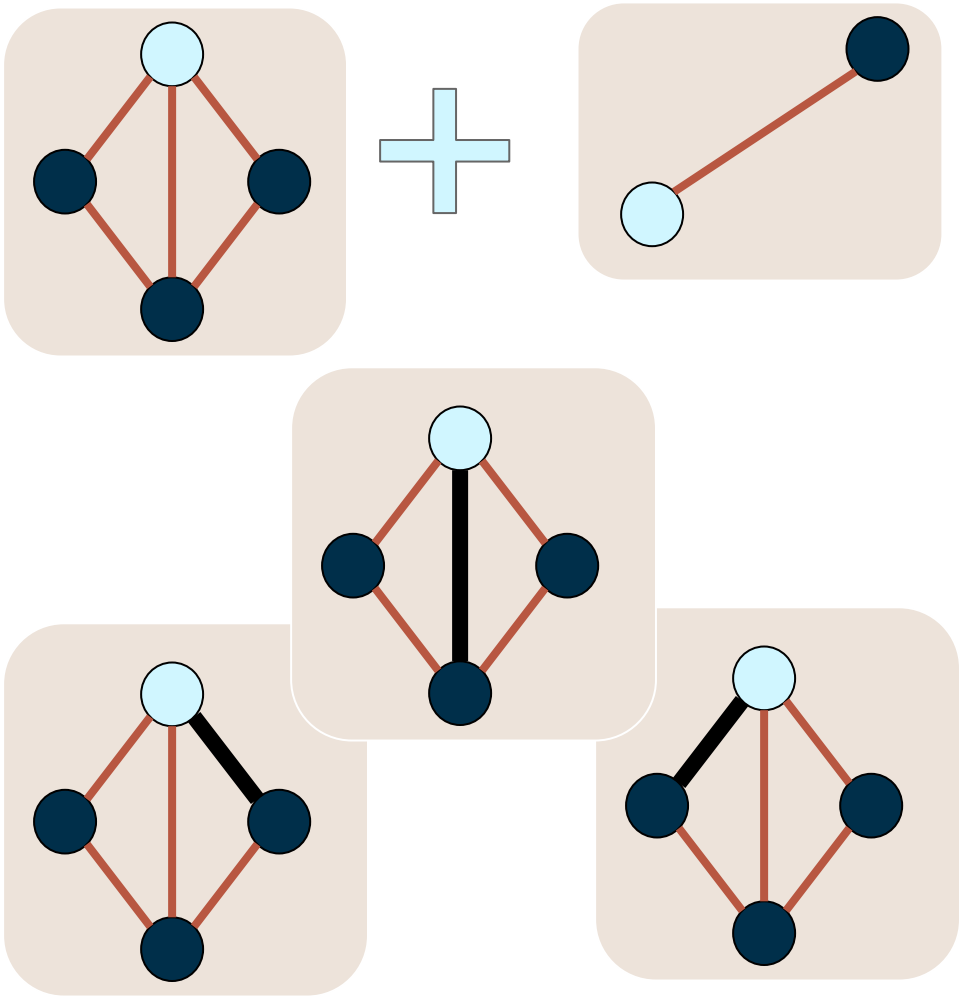


Fragments

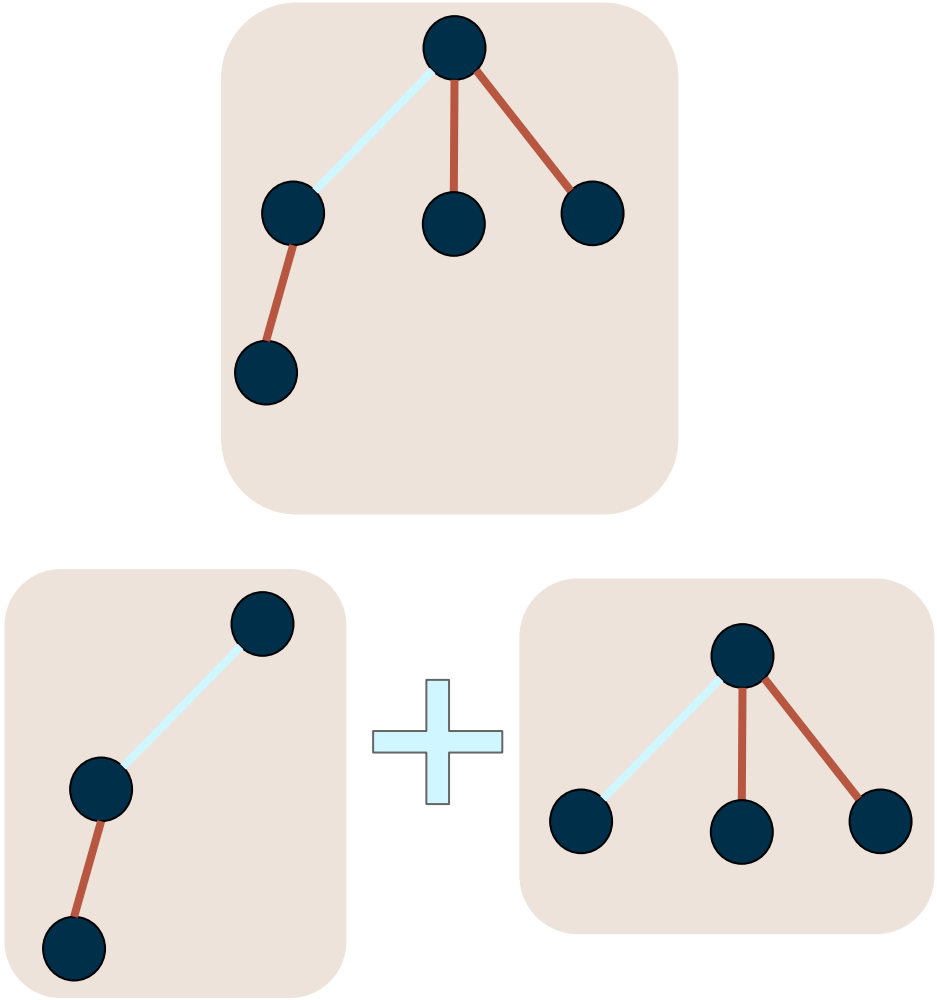




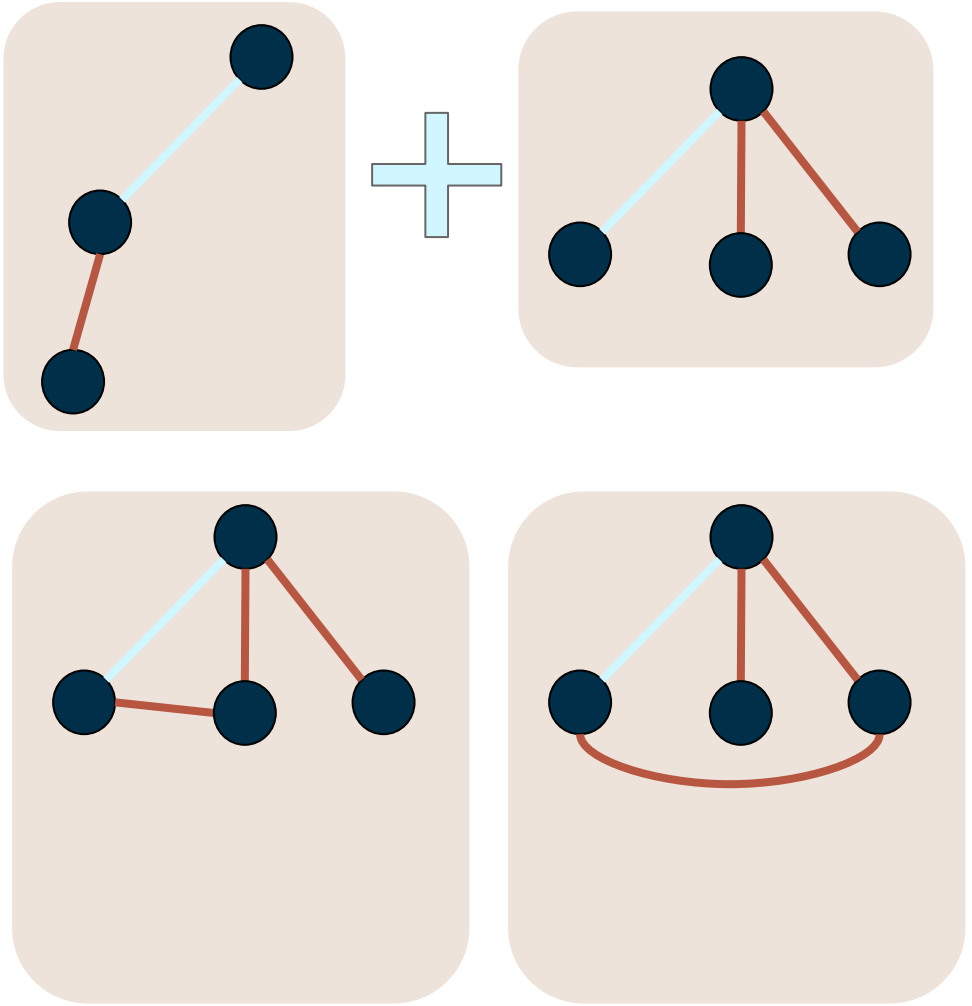
Shrinkage



Fragments



Shrinkage



Algorithm

Main Lemma

LEMMA 4. *Consider pattern H with cut set C. Then,*

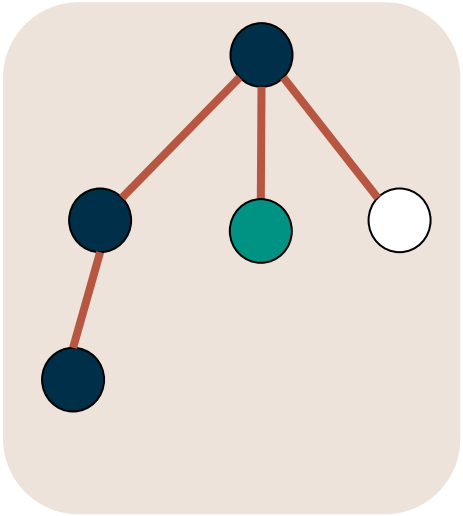
$$match(H) = \sum_{\sigma \in match(H|_C)} \prod_{F \in Frag_C(H)} \deg_F(\sigma) - \sum_{H' \in Shrink_C(H)} numSh_C(H, H') match(H')$$

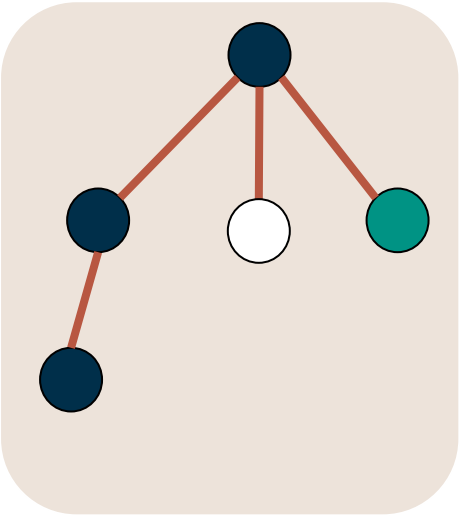
Algorithm

- Pattern count in all graph = sum over all the possible cuts count of pattern - total number of shortage.

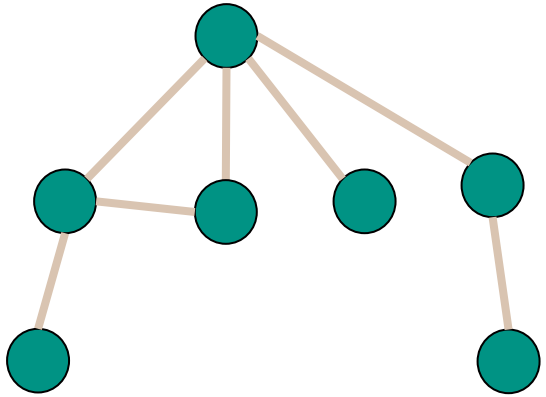
Algorithm

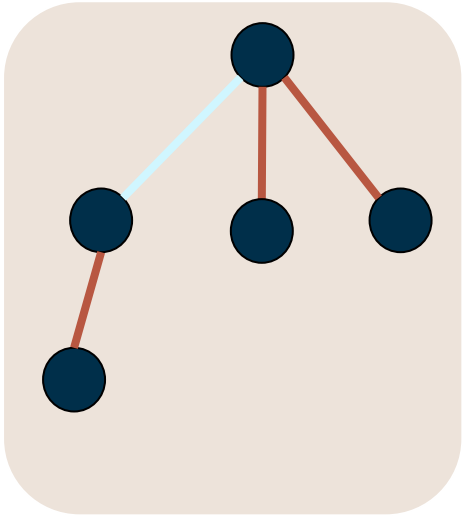
- Account for automorphisms



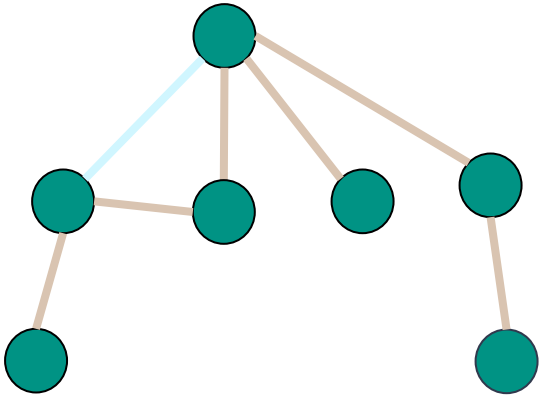


Example

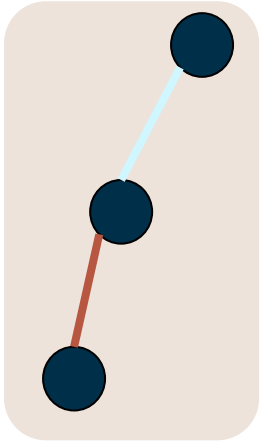




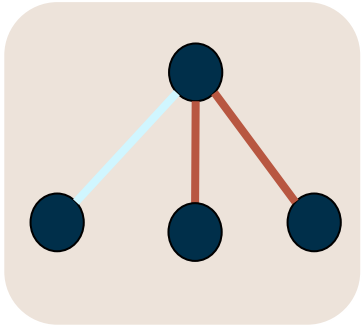
Example



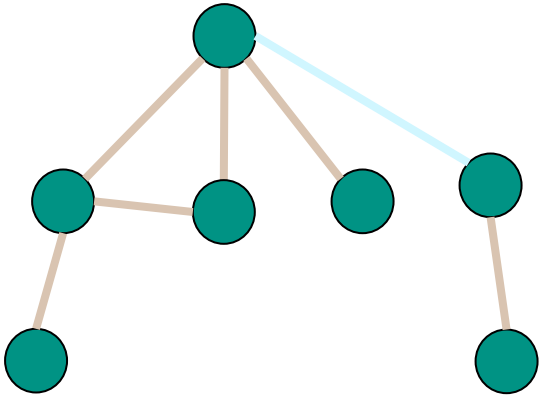
2X



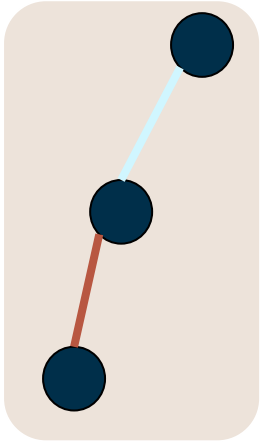
6X



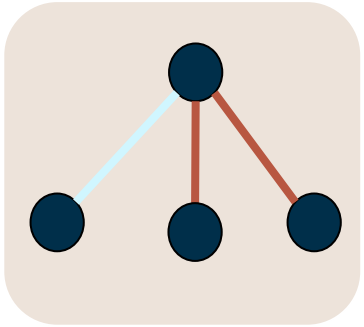
Example



1X



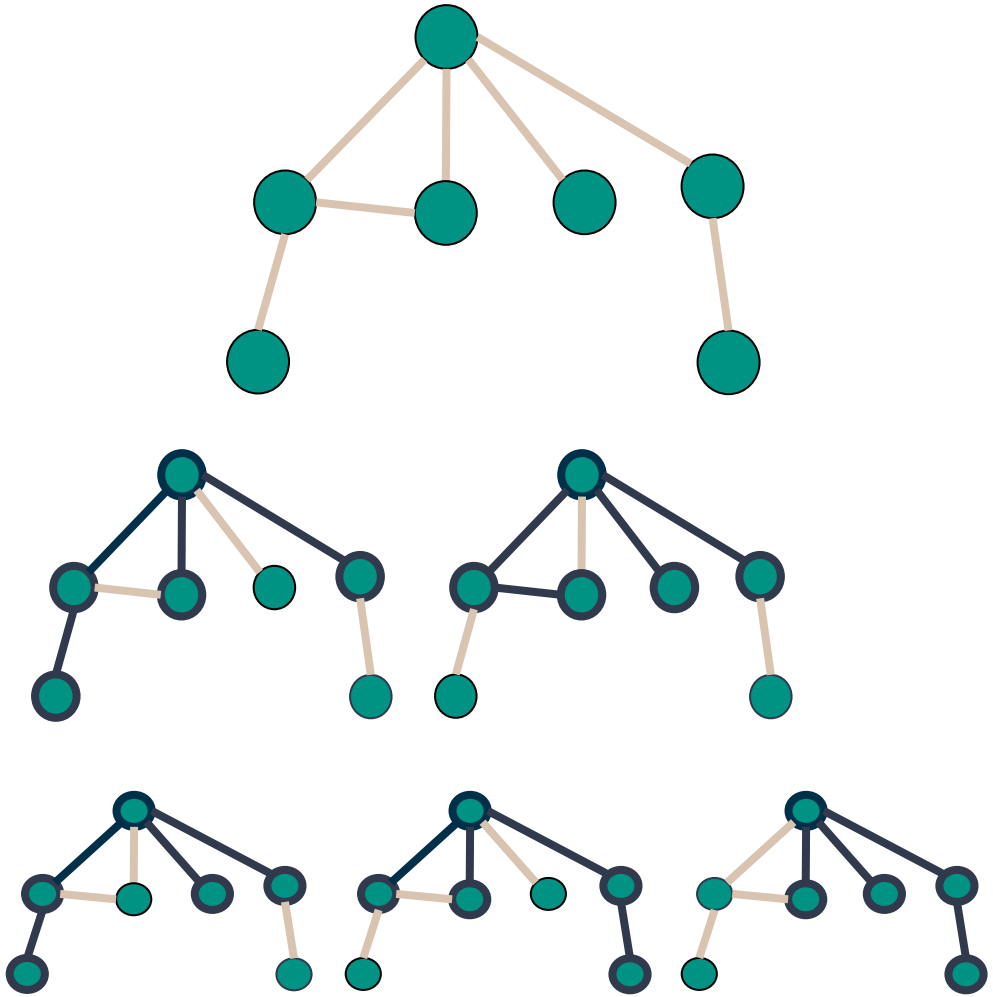
6X



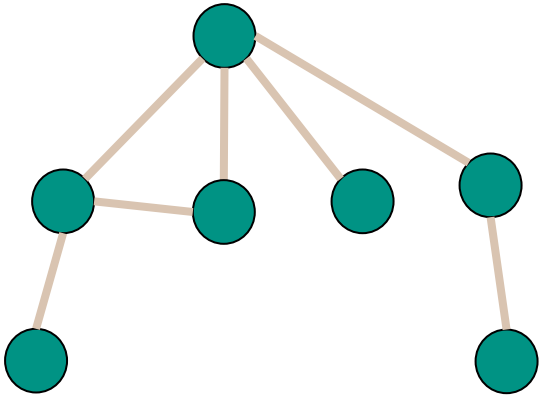
Algorithm

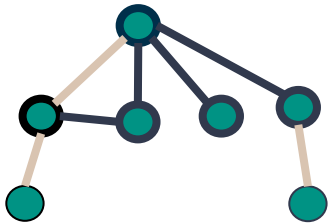
- Pattern count = 2*6 +1*6 - shrinkage
- Shrinkage = atomorphism * occurence = 2 * 2
- Answer = 18 - 4 = 14
- Accounting for automorphism = 14/2 = 7

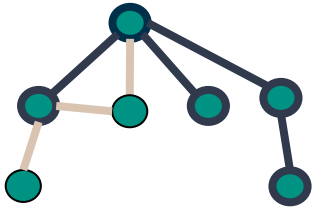
Example



Example







Results

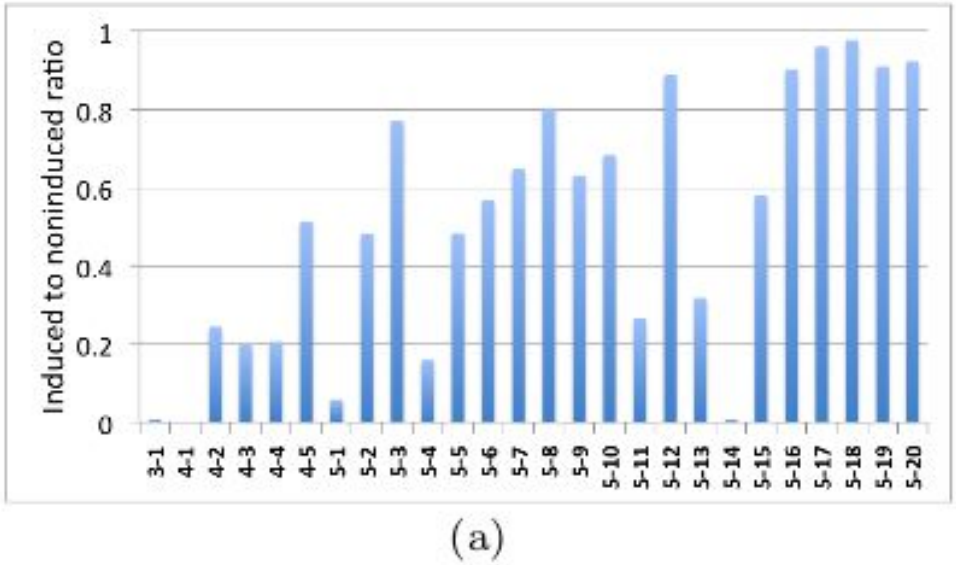
Performance

Comparison only with 4-vertex PGD

	$\|V\|$	$\|E\|$	$\|T\|$	Runtimes in seconds PGD	ESC-4	ESC-5
soc-brightkite	56.7K	426K	494K	1.20	0.22	6.54
tech-RL-caida	191K	1.22M	455K	3.21	0.25	5.47
flickr	244K	3.64M	15.9M	809K	12.9	961K
ia-email-EU-dir	265K	729K	267K	10.6	0.18	8.69
ca-coauth-dblp	540K	3.05M	444M	585	615	47.4K
web-google-dir	876K	8.64M	13.4M	54.5	2.94	71.8
tech-as-skitter	1,69M	22.2M	28.8M	1.90K	20.3	1.41K
web-wiki-ch-int	1.93M	9.16M	2.63M	4.91K	6.80	798
web-hudong	1.98M	14.6M	5.07M	9.40K	13.6	534
wiki-user-edits	2.09M	11.1M	6.68M	439K	2.92	9.15K
web-baidu-baike	2.14M	17.4M	3.57M	22.9K	16.2	9.46K
tech-ip	2.25M	21.6M	298K	613K	25.7	295
orkut	3.07M	234M	628M	598K	1.19K	217K
LiveJournal	4.84M	85.7M	286M	25.9K	538	37.1K

Edge Prediction

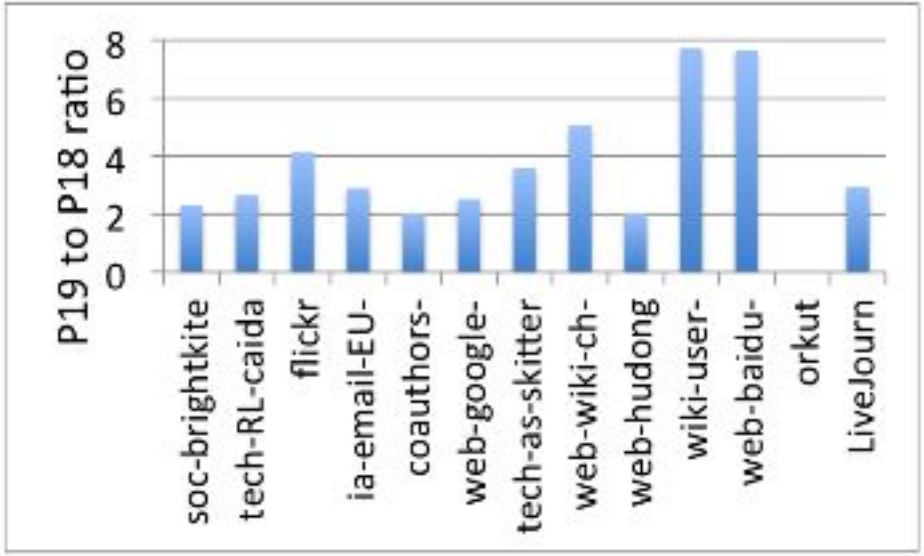
Prob. for having another edge



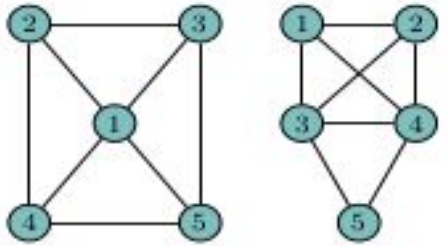
(a)

Subgraph Prediction

Ratio between patterns



(c)



(18) (19)

Future Work

- Scaling to 6-vertex subgraphs??!

References

- Pinar, Ali, C. Seshadhri, and Vaidyanathan Vishal. "Escape: Efficiently counting all 5-vertex subgraphs." In Proceedings of the 26th International Conference on World Wide Web, pp. 1431-1440. International World Wide Web Conferences Steering Committee, 2017.