

## COMPARISON B/W VARIOUS FP\_MINING TECHNIQUES

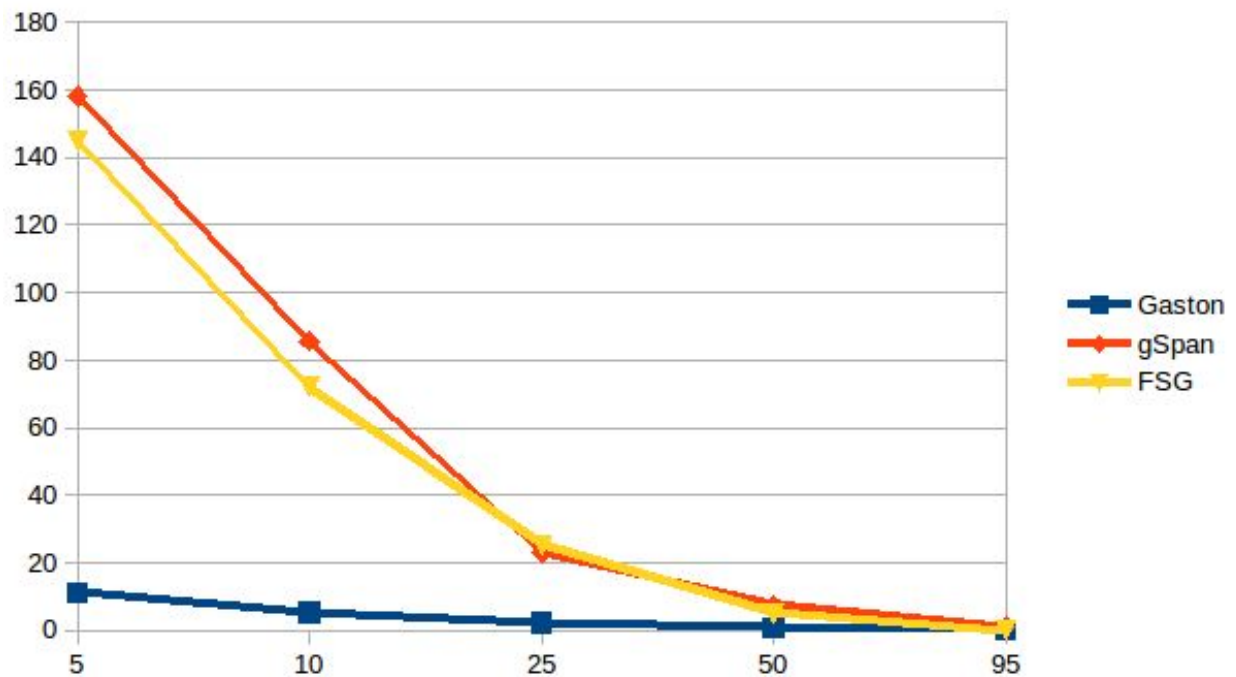
Implementations used:

GASTON- <http://liacs.leidenuniv.nl/~nijssensgr/gaston/index.html>

gSpan- <https://www.cs.ucsb.edu/~xyan/software/gSpan.htm>

PAFI/FSG- <http://glaros.dtc.umn.edu/gkhome/pafi/overview>

Support Ratio	Gaston	gSpan	FSG
5	11.22	158.252	144.7
10	5.53	85.61	72.3
25	2.4	23.15	25.3
50	0.91	7.67	5.2
95	0.29	1.39	0.1



It is visible from the graph that the running time of Gspan FSG increases exponentially on reducing minimum support value while for Gaston, running time is close to linear as compared with gSpan or FSG.

Here we see that Running time of g-Span is comparable to FSG and even higher for larger support values. This is unexpected as g-Span is usually faster because creation of size  $K+1$  candidate subgraphs from size  $K$  frequent Subgraphs is more complicated and costly than the standard Apriori Large itemset generation.

Also Pruning False positives is an expensive process.

FSG model used for the comparison has been developed well competitively in comparison with gSpan model.

Gaston is the fastest of all three models, although for support = 10, FSG evaluates subgraphs faster than FSG. This is because the Gaston takes advantage of the most frequent substructures in practical databases are free trees and uses this observation to implement an algorithm to boost its search by using a highly efficient enumeration strategy to enumerate free tree first.