

1- Setup

Attribute usage matrix for

Schedule (Pid, Aid, Date, Time, Route, NumPassenger)

will only do this for att's other than those in PK,
Since PK will be part of both fragments.

	A ₁	A ₂	A ₃	A ₄
q ₁		1	1	
q ₂	1			1
q ₃	1		1	

2- Access frequency

	S ₁	S ₂	S ₃	S ₄
q ₁	10	6	0	0
q ₂	0	0	8	10
q ₃	0	0	2	0

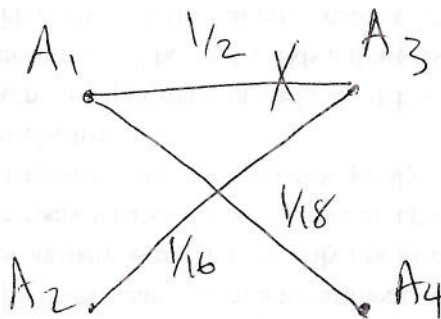
3 - Build att. affinity Matrix

$$\text{aff}(A_i, A_j) = \sum_{\substack{k | \text{use}(q_k, A_j) = 1 \wedge \\ \text{use}(q_k, A_i) = 1}} \sum_{\forall s_l} \text{acc}(q_k, s_l)$$

	A ₁	A ₂	A ₃	A ₄
A ₁	x*	0	2	18
A ₂		x*	16	0
A ₃			x*	0
A ₄				x*

x* diagonal terms
don't play a roll
in this method
so just left them out.

4 - Build graph showing distance between att's



5 - MST is the same as the graph itself

A₁, A₃ is the highest cost edge on MST, so we remove it
to find following two clusters/fragments (A₁, A₄), (A₂, A₃)

6 - The two vertical fragments will be (Pid, Aid, Date, NumPers) & (Pid, Aid, Time, ...)