Answer

The upper bound of threshold ɑ(ɑ\_max) = min(U), which means ɑ\_max is equal to the minimal Upper Bound among all term. For example, the ɑ\_max is 3 for below table.

|  |  |  |  |
| --- | --- | --- | --- |
| term | a | b | c |
| Upper bound | 4 | 3 | 5 |

Prove:

As long as ɑ\_max = min(U), then ɑ < ɑ\_max, every potential candidate can be calculated, which mean we won’t omit any candidates.

But if ɑ >= ɑ\_max, we may miss some candidates that can be true candidate. For example, if we have two term A,B, k = 3 and their inverted\_index is shown as below table:

|  |  |
| --- | --- |
| A | <1,3>, <2,1>, <3,8>, <4,4> |
| B | <1,4>, <2,2>, <3,9>, <5,1> |

The answer for this condition should be [(7,1), (17,3), (4,4)], if we set ɑ = min(U) = 8, then it would lead to a wrong answer [(7,1), (3,2), (17,3)] as <4,4> can’t be evaluated as the upper bound of A is 8 which is not bigger than ɑ.