

Aim: To implement the HITS algorithm.

Theory: The (Hyperlink Induced Topic Search) HITS algorithm also known as the hubs and authorities algorithm is a link analysis algorithm used in data mining & information retrieval. It was designed to evaluate importance of web pages in hyperlinked environment.

Brief overview of the HITS algorithm.

- 1) Hubs & authorities: A hub is a page that links to many other pages. Authority is a page that is linked by many hubs.
- 2) Algorithm iterations: The algorithm operates through iterative updates where the authority & hub scores its pages over refining in each iteration.
- 3) Score calculation: Initially all the pages are given equal hub & authority scores.
- 4) Normalisation - After each iteration the scores are normalised.
- 5) Convergence - The ~~iter~~ algorithm continues to iterate until the hub & authority scores stabilize indicating convergence.

Conclusion: Thus we successfully implemented the HITS algorithm.