

# HITS ALGORITHM

SEMESTER PROJECT REPORT ANALYSIS AND  
ALGORITHM



<b>Submitted by :</b>	Basit Ali
<b>Submitted to:</b>	Mr. Usman Sharif
<b>SAP:</b>	54596
<b>Section:</b>	CS 4-1

**DATE:** May 11, 2025

# Web Page Ranking Using HITS Algorithm in C++

---

## 1. Introduction:

The Hyperlink-Induced Topic Search (HITS) algorithm, also known as Hubs and Authorities, is a link analysis algorithm used to rank web pages. It was developed by Jon Kleinberg. This algorithm identifies two types of web pages:

- Authorities: Pages that contain useful information.
- Hubs: Pages that link to many authorities.

## 2. Problem Statement:

Let suppose we have a set of web pages and the links between them, identify the most authoritative pages and the best hub pages using the HITS algorithm implemented in C++.

## 3. Methodology:

- Accept the number of nodes (web pages) and their link structure using an adjacency matrix.
- Initialize all hub and authority scores to 1.
- Update authority scores based on incoming links (linked from hubs).
- Update hub scores based on outgoing links (linked to authorities).
- Normalize the scores to prevent overflow.
- Repeat the process for a fixed number of iterations (e.g., 5).
- Output the final authority and hub scores.

## 4. Code Explanation:

- It is a C++ programming language code
- Takes input for the number of web pages and accepts an adjacency matrix (using 1 for a link and 0 for no link), then initializes hub and authority scores to 1 or 0.
- In each iteration it will calculate new authority scores using current hub scores and
- calculates new hub scores using current authority scores.
- It normalizes both vectors and after several iterations, it displays the final normalized score

## 5. Time Complexity Analysis:

- Let  $N$  be the number of nodes: Reading input:  $O(N^2)$
- Each iteration: Authority and hub update:  $O(N^2)$
- Normalization:  $O(N)$

So, Total for  $k$  iterations (fixed, say 5):  $O(k \times N^2)$

## 6. Sample Input and Output

- Input:

Enter the number of nodes: 3

Enter the adjacency matrix:

0 1 1

0 0 1

1 0 0

- Output:

Final Authority and Hub Scores:

Node 0: Authority = 0.5161, Hub = 0.6110

Node 1: Authority = 0.6804, Hub = 0.1393

Node 2: Authority = 0.5193, Hub = 0.7784

## 7. Conclusion:

The HITS algorithm is useful in analyzing directed graphs to identify important nodes. It separates nodes into hubs and authorities, which can be used to rank search results or detect influence in networks. The C++ implementation helps beginners understand how iterative link analysis works using basic arrays, loops, and mathematical normalization.

**GITHUB LINK :-**

<https://github.com/BasitAli9/Project-of-Algorithm-and-analysis.git>