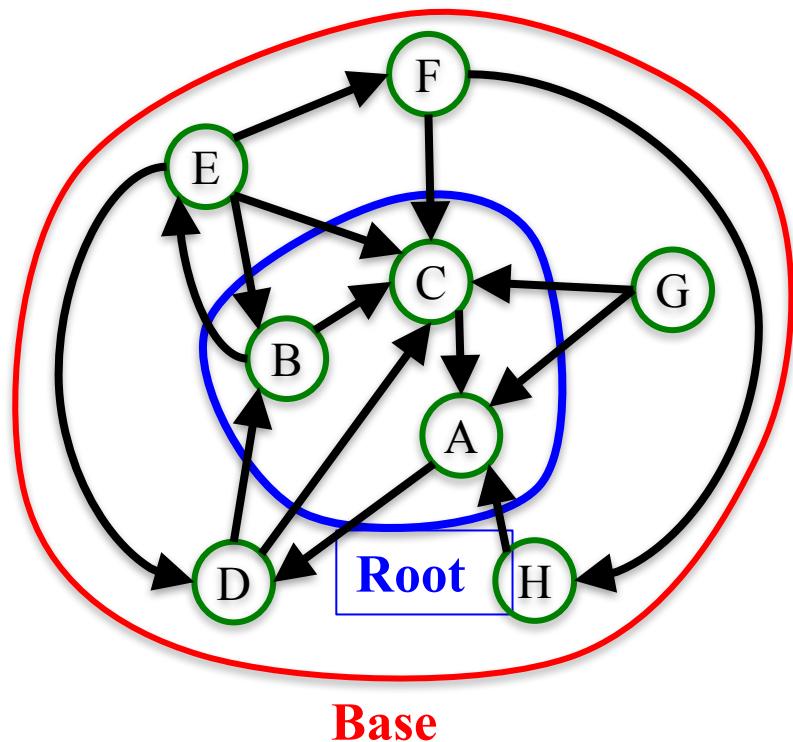


Hubs and Authorities

Given a query to a search engine:

- **Root**: set of highly relevant web pages (e.g. pages that contain the query string) – potential *authorities*.
- Find all pages that link to a page in root – potential *hubs*.
- **Base**: root nodes and any node that links to a node in root.
- Consider all edges connecting nodes in the base set.



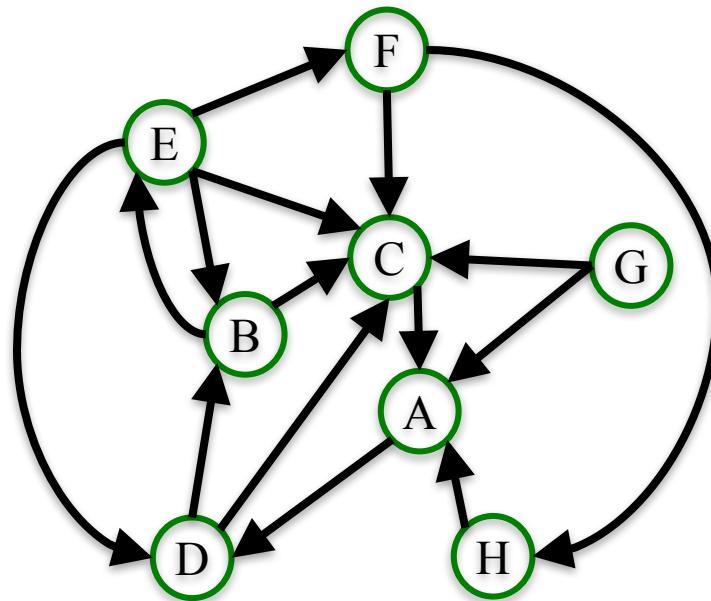
HITS Algorithm

Computing k iterations of the HITS algorithm to assign an *authority score* and *hub score* to each node.

1. Assign each node an authority and hub score of 1.
2. Apply the **Authority Update Rule**: each node's *authority* score is the sum of *hub* scores of each node that *points to it*.
3. Apply the **Hub Update Rule**: each node's *hub* score is the sum of *authority* scores of each node that *it points to*.
4. **Normalize** Authority and Hub scores: $\text{auth}(j) = \frac{\text{auth}(j)}{\sum_{i \in N} \text{auth}(i)}$
5. Repeat k times.

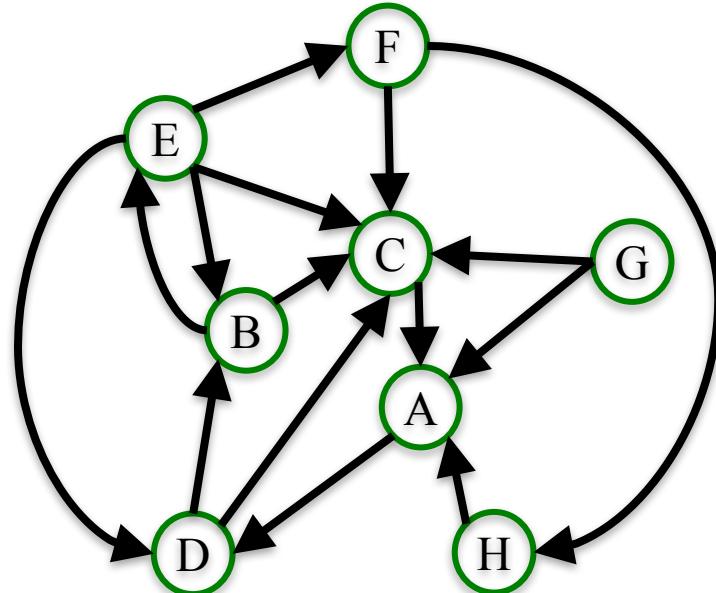
HITS Algorithm Example

Compute 2 iterations of the HITS algorithm on this network.



HITS Algorithm Example

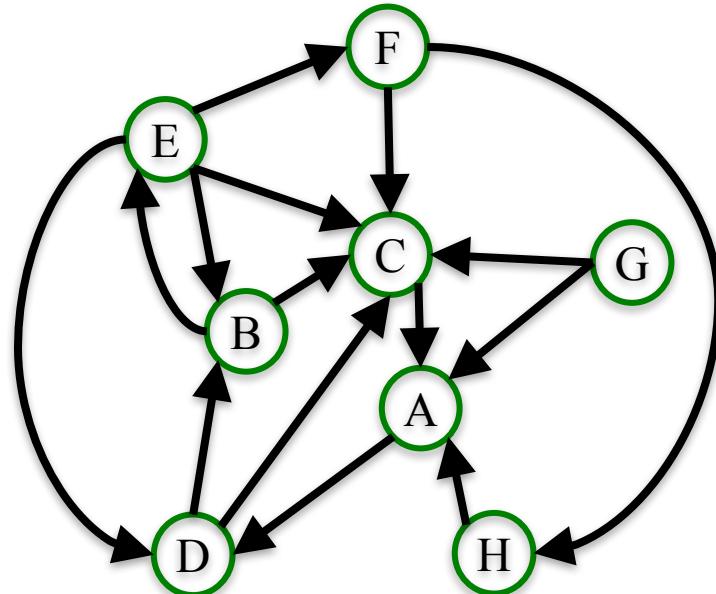
	Old Auth	Old Hub	New Auth	New Hub
A				
B				
C				
D				
E				
F				
G				
H				



Assign each node an auth and hub score of 1

HITS Algorithm Example

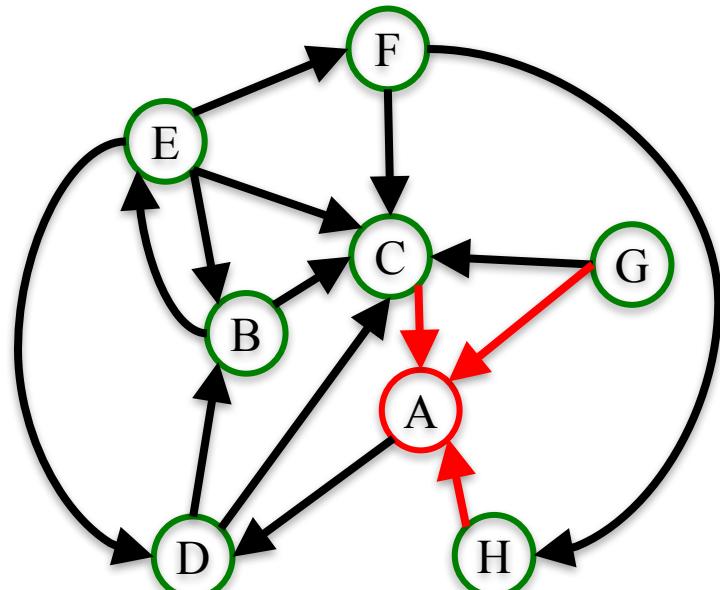
	Old Auth	Old Hub	New Auth	New Hub
A	1	1		
B	1	1		
C	1	1		
D	1	1		
E	1	1		
F	1	1		
G	1	1		
H	1	1		



Assign each node an auth and hub score of 1

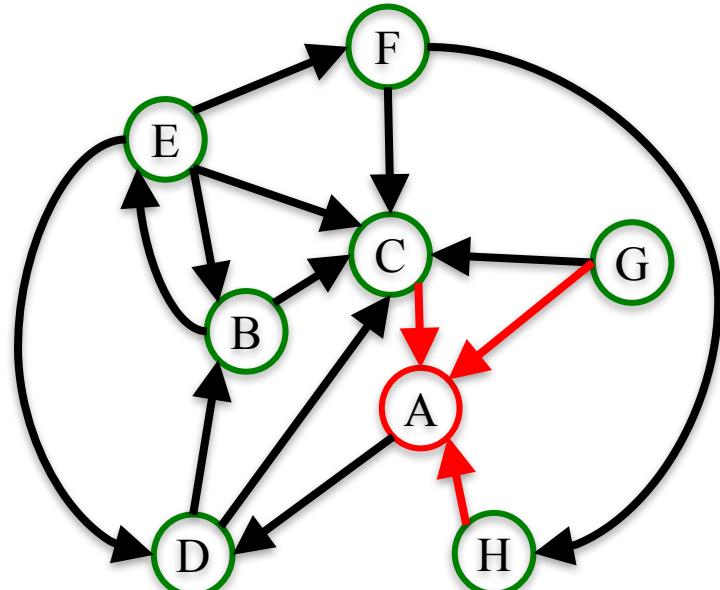
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1		
B	1	1		
C	1	1		
D	1	1		
E	1	1		
F	1	1		
G	1	1		
H	1	1		



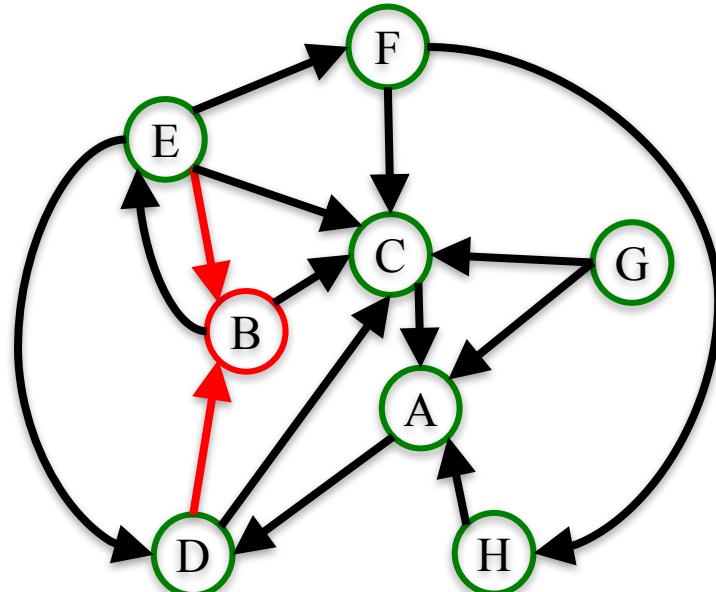
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1		
C	1	1		
D	1	1		
E	1	1		
F	1	1		
G	1	1		
H	1	1		



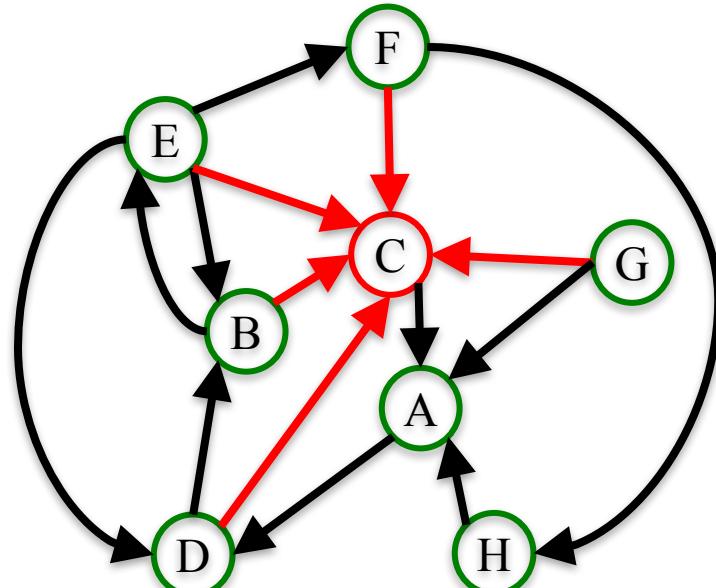
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1	2	
C	1	1		
D	1	1		
E	1	1		
F	1	1		
G	1	1		
H	1	1		



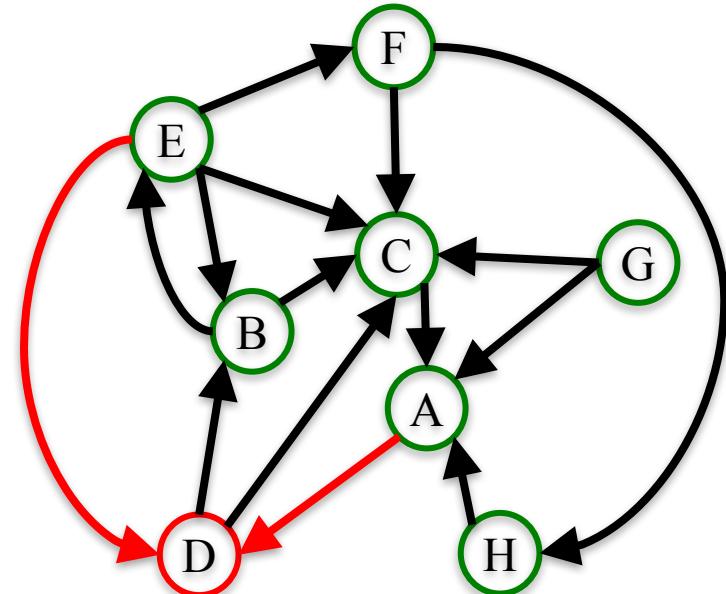
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1	2	
C	1	1	5	
D	1	1		
E	1	1		
F	1	1		
G	1	1		
H	1	1		



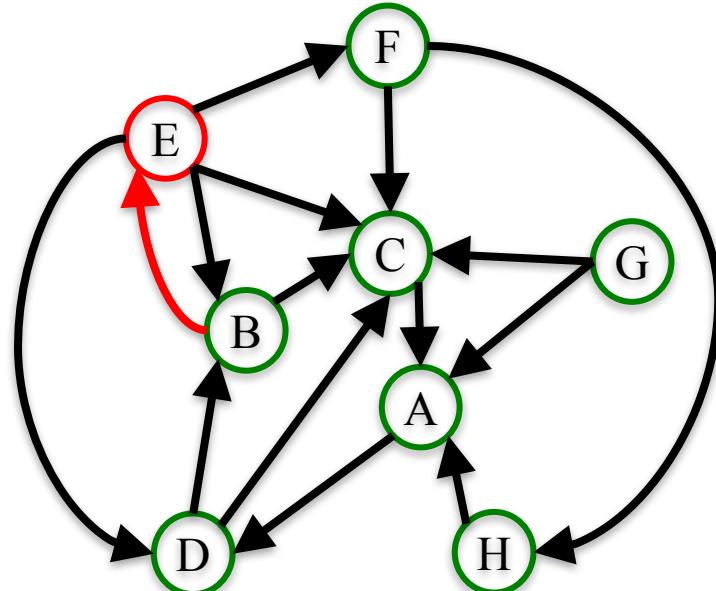
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1	2	
C	1	1	5	
D	1	1	2	
E	1	1		
F	1	1		
G	1	1		
H	1	1		



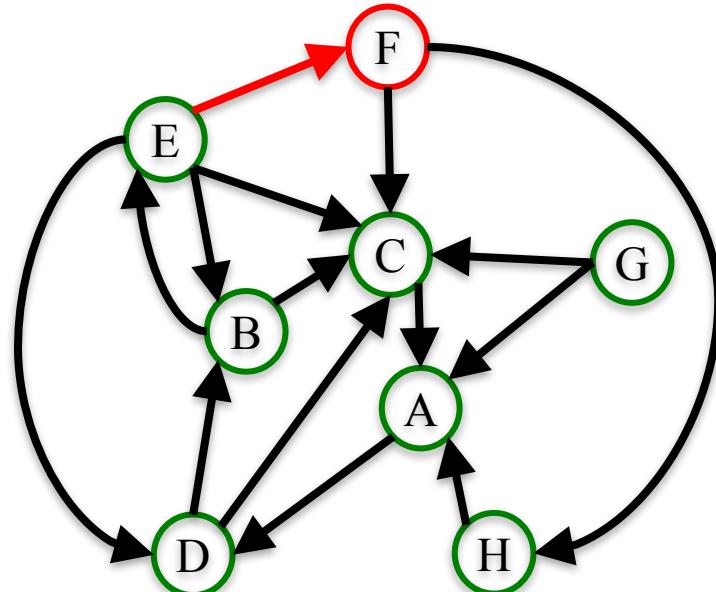
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1	2	
C	1	1	5	
D	1	1	2	
E	1	1	1	
F	1	1		
G	1	1		
H	1	1		



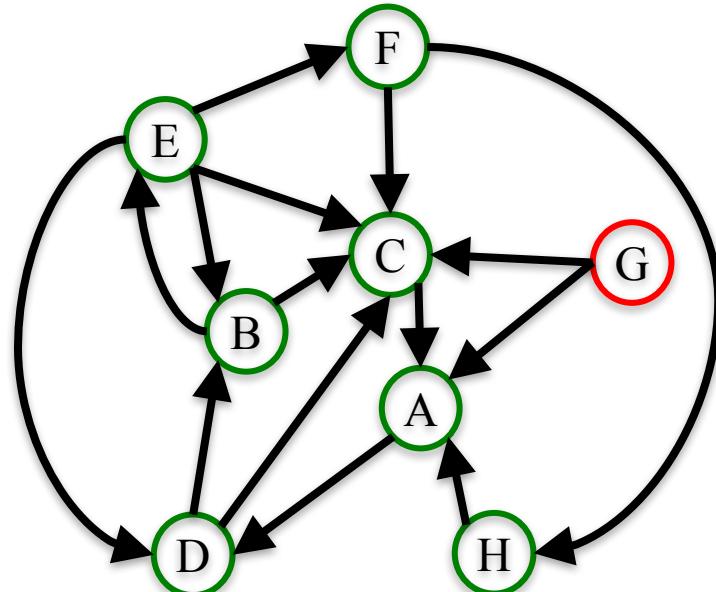
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1	2	
C	1	1	5	
D	1	1	2	
E	1	1	1	
F	1	1	1	
G	1	1		
H	1	1		



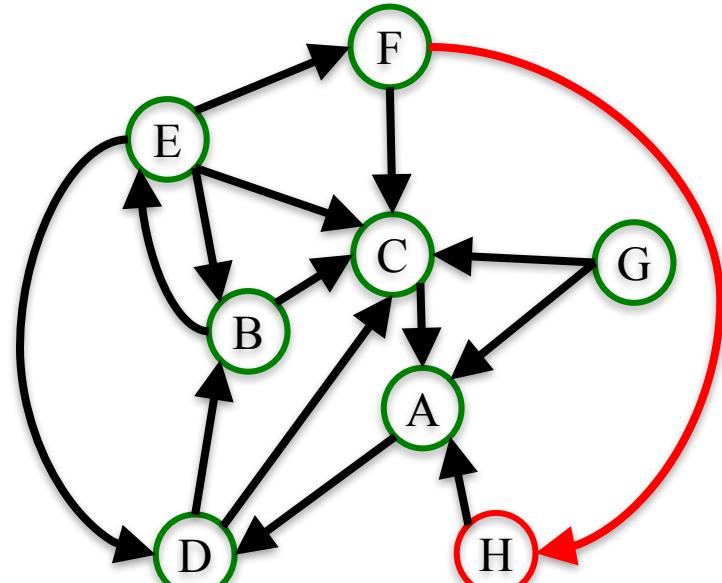
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1	2	
C	1	1	5	
D	1	1	2	
E	1	1	1	
F	1	1	1	
G	1	1	0	
H	1	1		



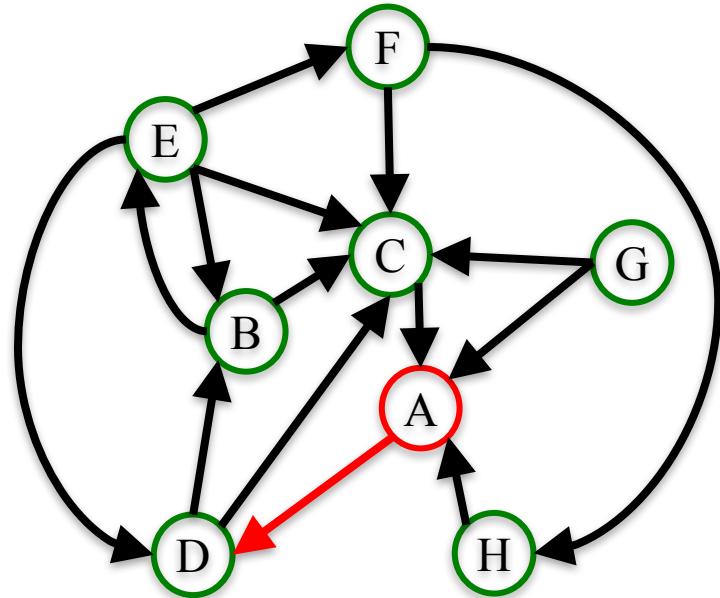
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1	2	
C	1	1	5	
D	1	1	2	
E	1	1	1	
F	1	1	1	
G	1	1	0	
H	1	1	1	



HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	
B	1	1	2	
C	1	1	5	
D	1	1	2	
E	1	1	1	
F	1	1	1	
G	1	1	0	
H	1	1	1	

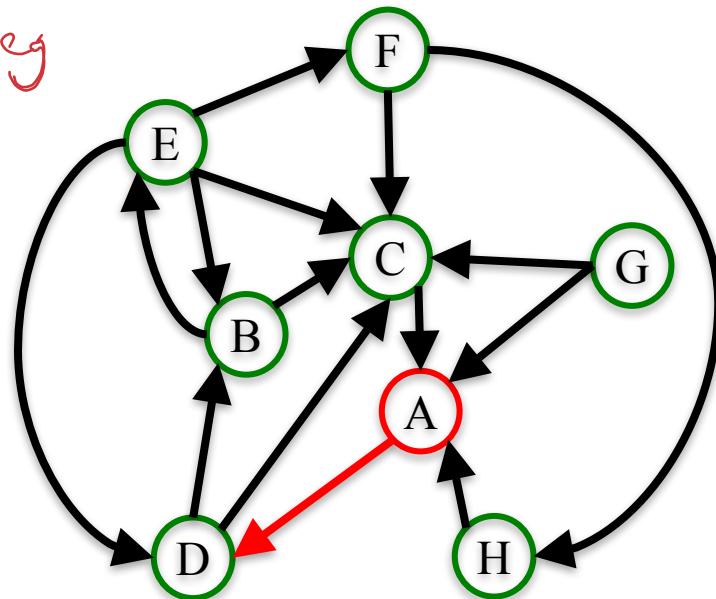


in-deg

HITS Algorithm Example

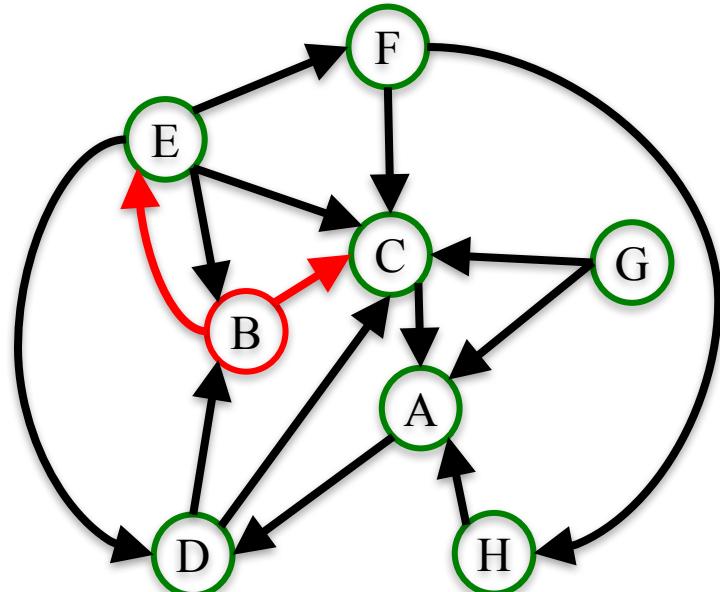
out-deg

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	
C	1	1	5	
D	1	1	2	
E	1	1	1	
F	1	1	1	
G	1	1	0	
H	1	1	1	



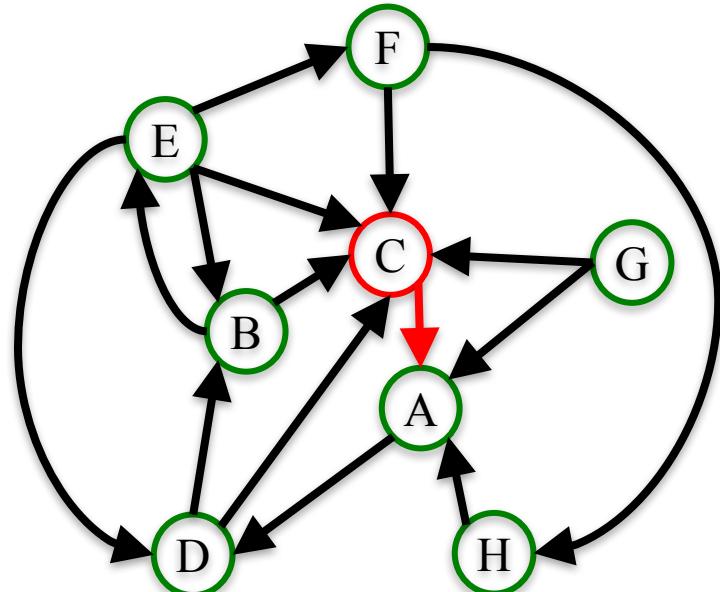
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	2
C	1	1	5	
D	1	1	2	
E	1	1	1	
F	1	1	1	
G	1	1	0	
H	1	1	1	



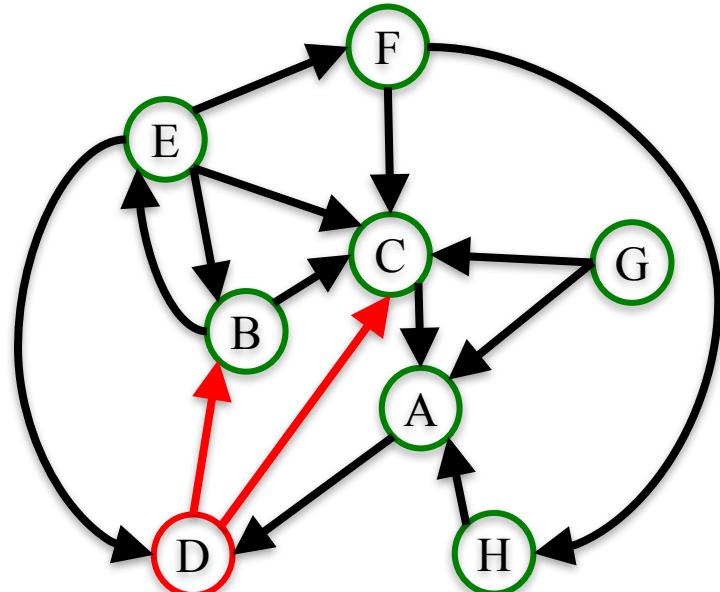
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	2
C	1	1	5	1
D	1	1	2	
E	1	1	1	
F	1	1	1	
G	1	1	0	
H	1	1	1	



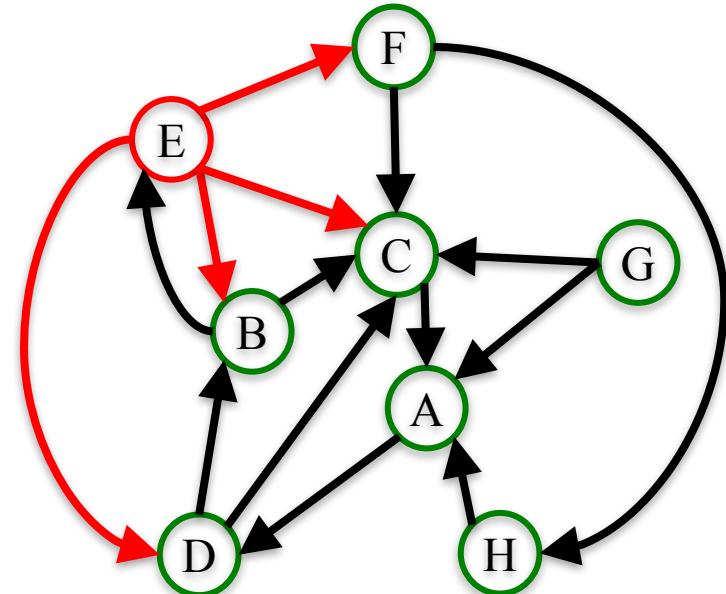
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	2
C	1	1	5	1
D	1	1	2	2
E	1	1	1	
F	1	1	1	
G	1	1	0	
H	1	1	1	



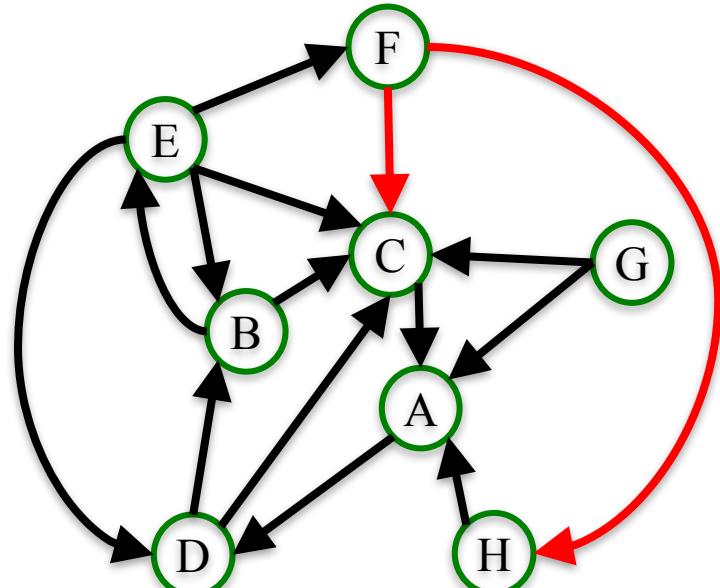
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	2
C	1	1	5	1
D	1	1	2	2
E	1	1	1	4
F	1	1	1	
G	1	1	0	
H	1	1	1	



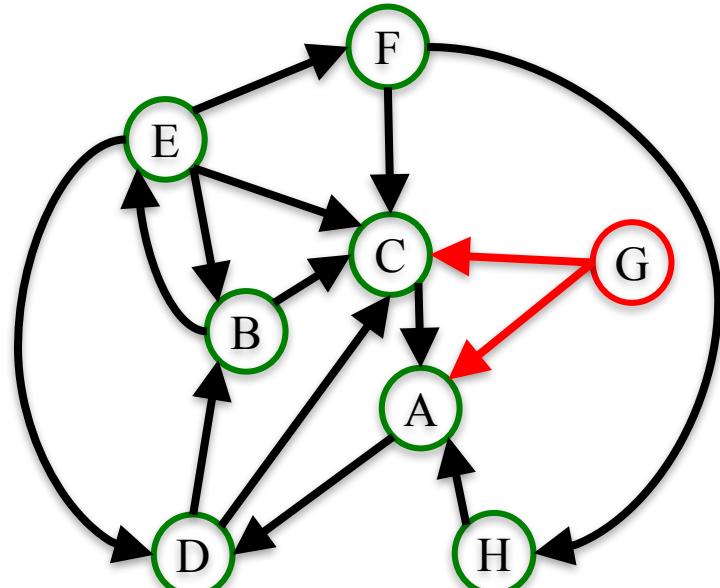
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	2
C	1	1	5	1
D	1	1	2	2
E	1	1	1	4
F	1	1	1	2
G	1	1	0	
H	1	1	1	



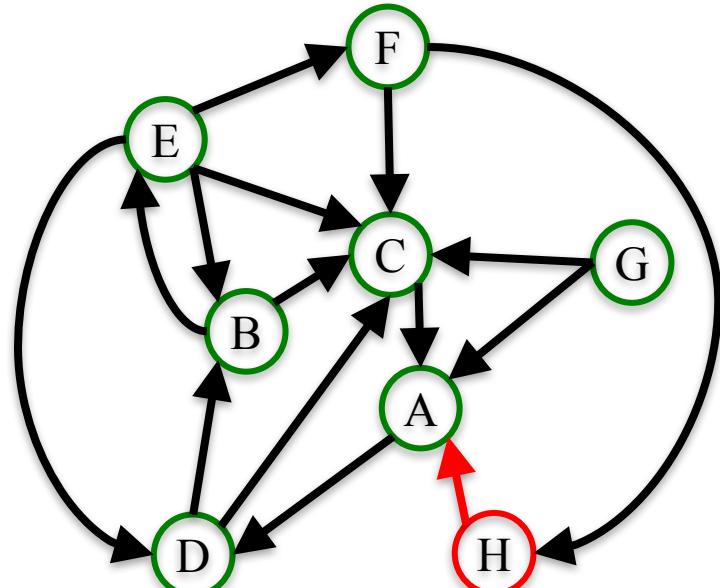
HITS Algorithm

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	2
C	1	1	5	1
D	1	1	2	2
E	1	1	1	4
F	1	1	1	2
G	1	1	0	2
H	1	1	1	



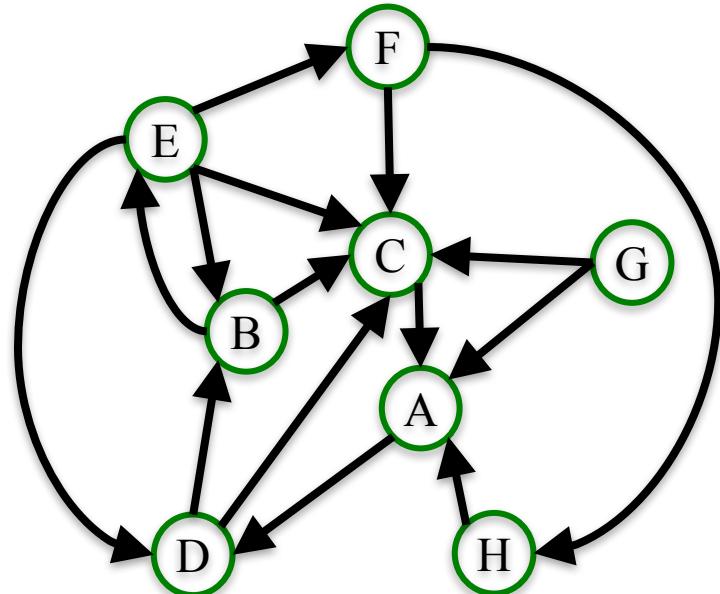
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	2
C	1	1	5	1
D	1	1	2	2
E	1	1	1	4
F	1	1	1	2
G	1	1	0	2
H	1	1	1	1



HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3	1
B	1	1	2	2
C	1	1	5	1
D	1	1	2	2
E	1	1	1	4
F	1	1	1	2
G	1	1	0	2
H	1	1	1	1

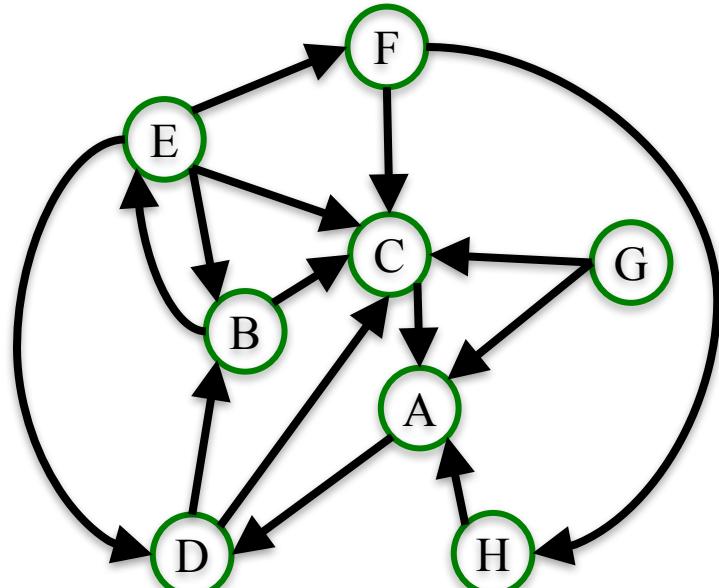


Normalize:

$$\sum_{i \in N} \text{auth}(i) = 15 \quad \sum_{i \in N} \text{hub}(i) = 15$$

HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1	1	3/15	1/15
B	1	1	2/15	2/15
C	1	1	5/15	1/15
D	1	1	2/15	2/15
E	1	1	1/15	4/15
F	1	1	1/15	2/15
G	1	1	0/15	2/15
H	1	1	1/15	1/15

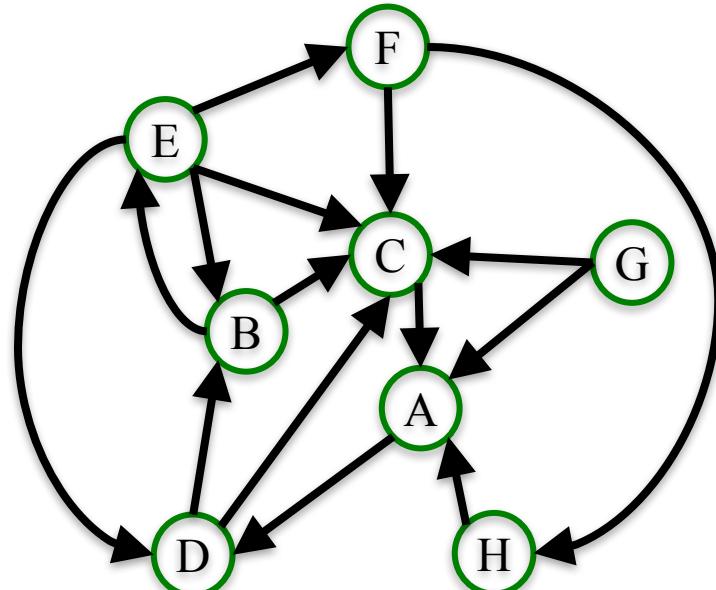


Normalize:

$$\sum_{i \in N} \text{auth}(i) = 15 \quad \sum_{i \in N} \text{hub}(i) = 15$$

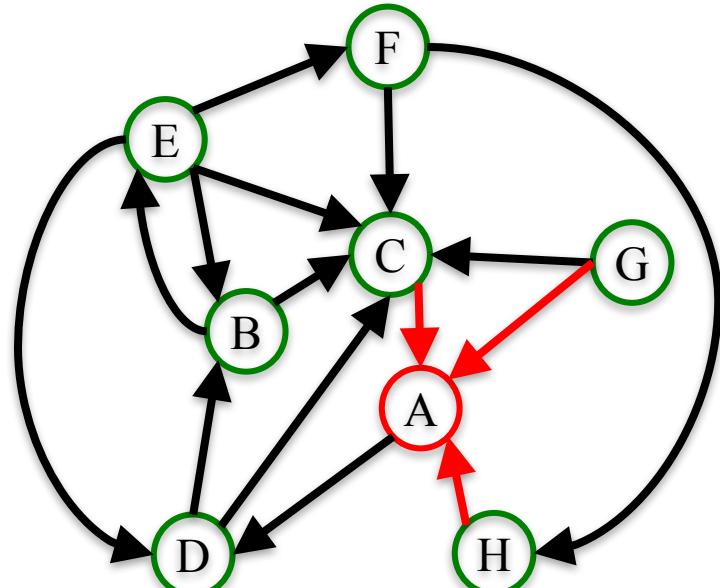
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15		
B	2/15	2/15		
C	1/3	1/15		
D	2/15	2/15		
E	1/15	4/15		
F	1/15	2/15		
G	0	2/15		
H	1/15	1/15		



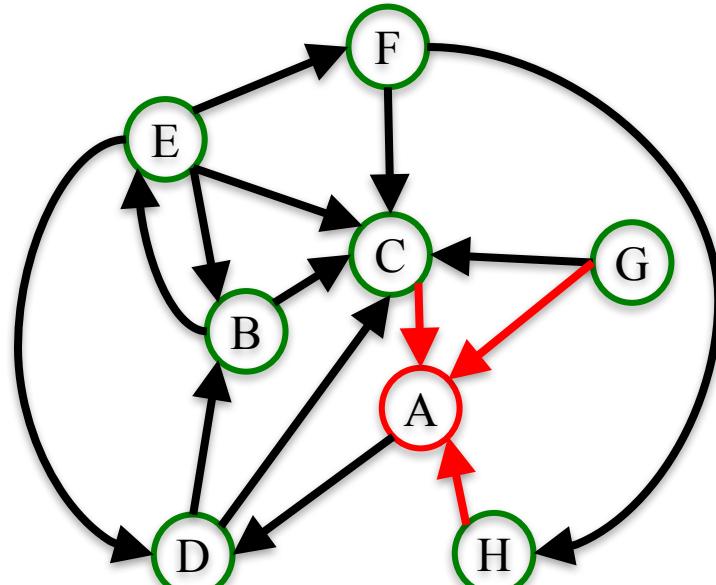
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15		
B	2/15	2/15		
C	1/3	1/15		
D	2/15	2/15		
E	1/15	4/15		
F	1/15	2/15		
G	0	2/15		
H	1/15	1/15		



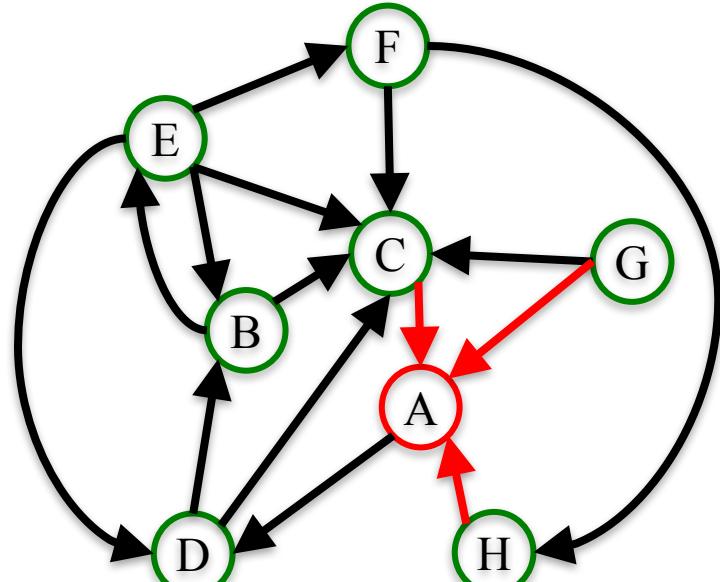
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15		
B	2/15	2/15		
C	1/3	1/15		
D	2/15	2/15		
E	1/15	4/15		
F	1/15	2/15		
G	0	2/15		
H	1/15	1/15		



HITS Algorithm Example

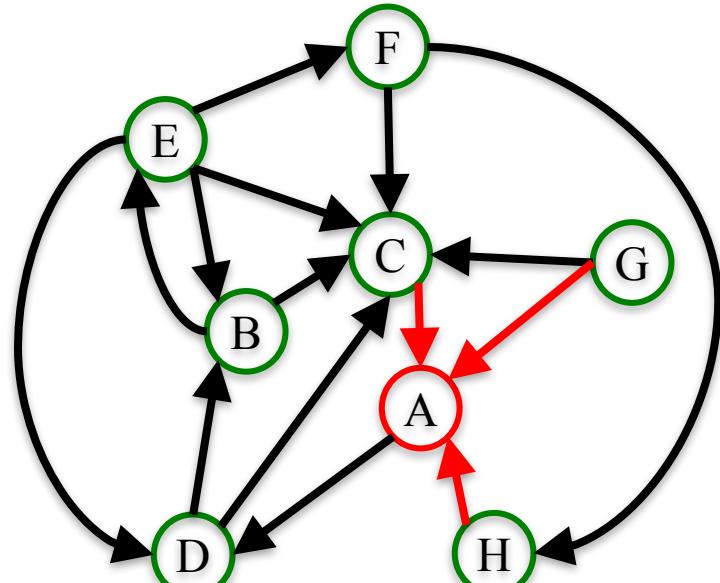
	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15		
B	2/15	2/15		
C	1/3	1/15		
D	2/15	2/15		
E	1/15	4/15		
F	1/15	2/15		
G	0	2/15		
H	1/15	1/15		



$$\frac{1}{15} + \frac{2}{15} + \frac{1}{15} = \frac{4}{15}$$

HITS Algorithm Example

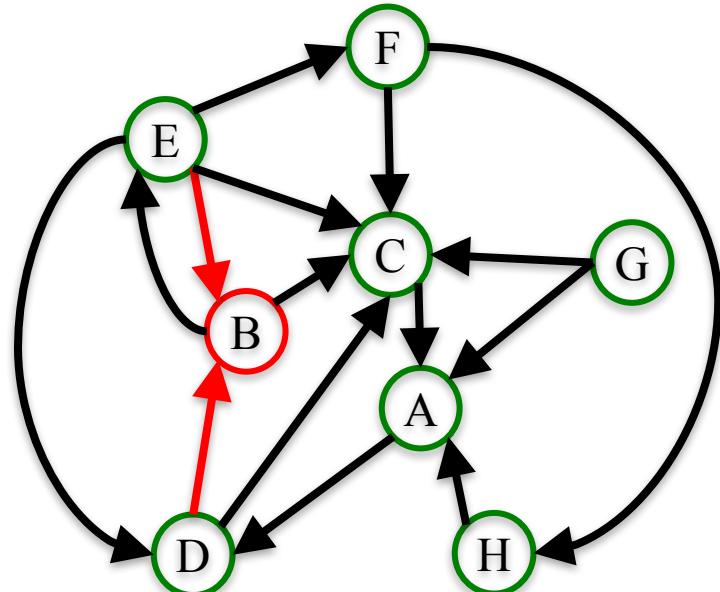
	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	
B	2/15	2/15		
C	1/3	1/15		
D	2/15	2/15		
E	1/15	4/15		
F	1/15	2/15		
G	0	2/15		
H	1/15	1/15		



$$\frac{1}{15} + \frac{2}{15} + \frac{1}{15} = \frac{4}{15}$$

HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	
B	2/15	2/15	6/15	
C	1/3	1/15		
D	2/15	2/15		
E	1/15	4/15		
F	1/15	2/15		
G	0	2/15		
H	1/15	1/15		

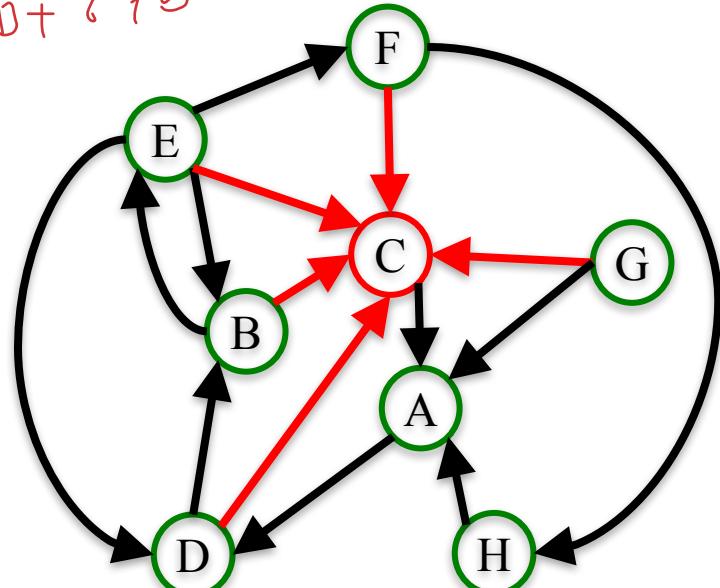


$$\frac{2}{15} + \frac{4}{15} = \frac{6}{15}$$

HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	
B	2/15	2/15	6/15	
C	1/3	1/15	12/15	
D	2/15	2/15		
E	1/15	4/15		
F	1/15	2/15		
G	0	2/15		
H	1/15	1/15		

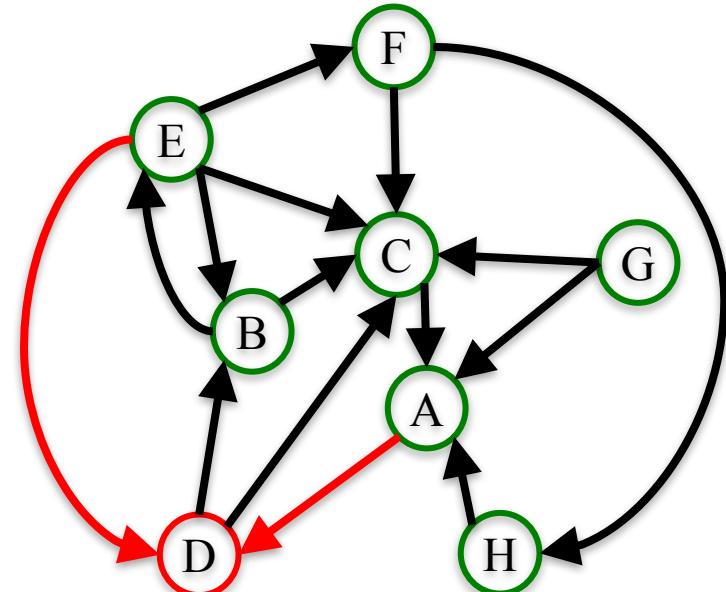
E + F + D + C = 15



$$\frac{2}{15} + \frac{2}{15} + \frac{4}{15} + \frac{2}{15} + \frac{2}{15} = \frac{12}{15}$$

HITS Algorithm Example

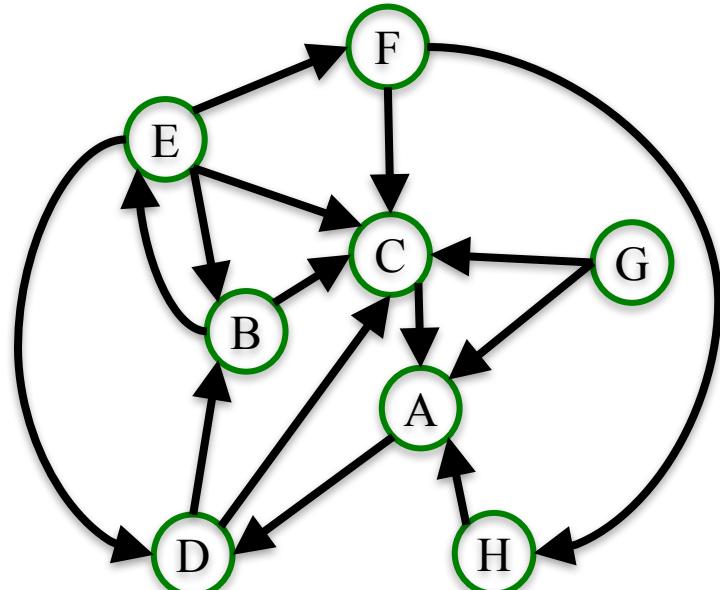
	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	
B	2/15	2/15	6/15	
C	1/3	1/15	12/15	
D	2/15	2/15	1/3	
E	1/15	4/15		
F	1/15	2/15		
G	0	2/15		
H	1/15	1/15		



$$\frac{1}{15} + \frac{4}{15} = \frac{5}{15} = \frac{1}{3}$$

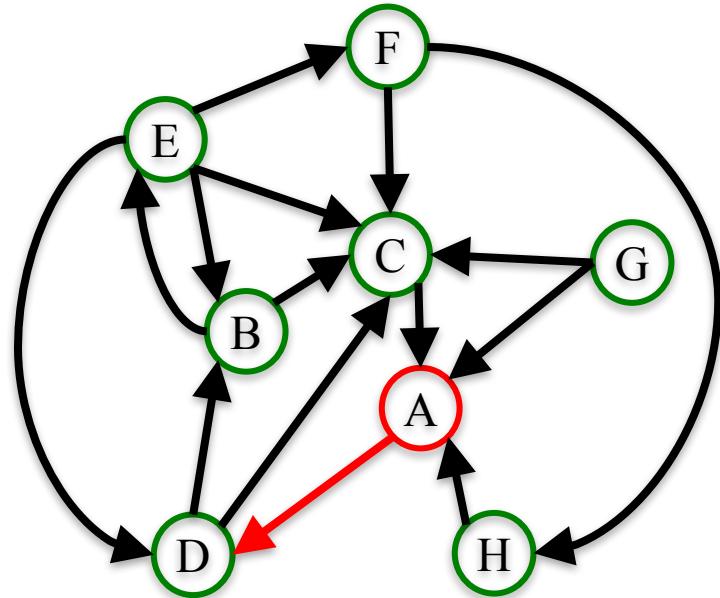
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	
B	2/15	2/15	6/15	
C	1/3	1/15	12/15	
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



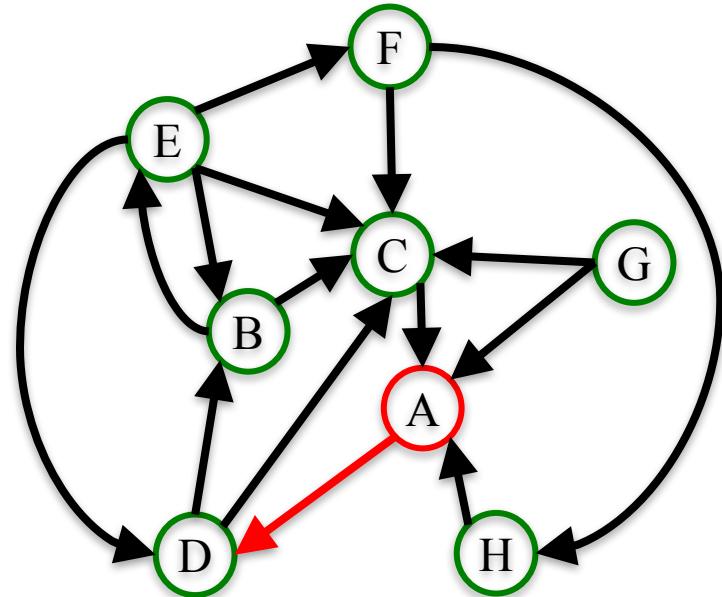
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	
B	2/15	2/15	6/15	
C	1/3	1/15	12/15	
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



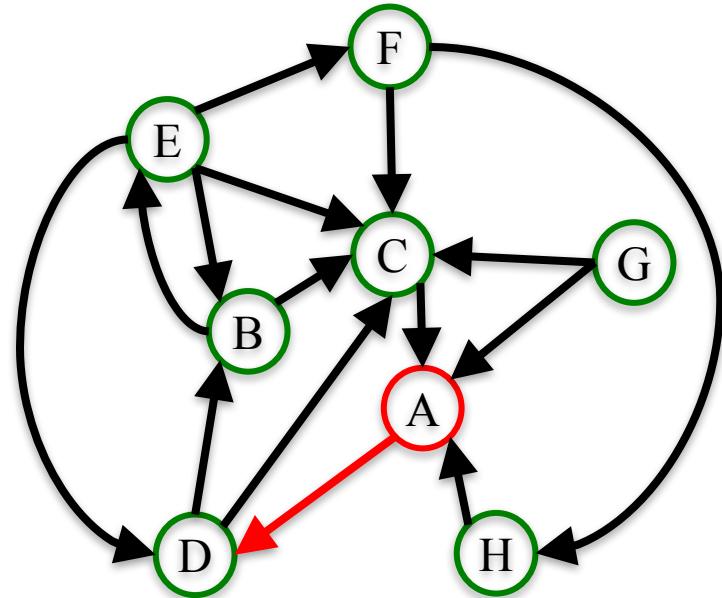
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	
B	2/15	2/15	6/15	
C	1/3	1/15	12/15	
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



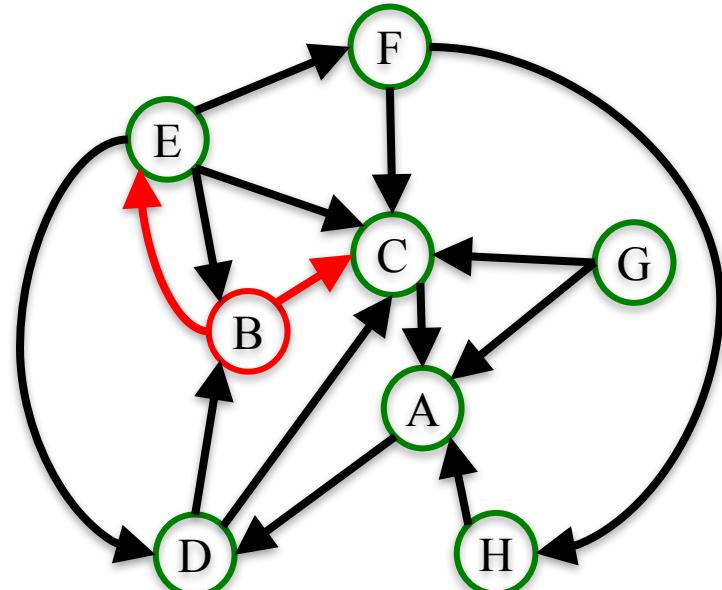
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	
C	1/3	1/15	12/15	
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



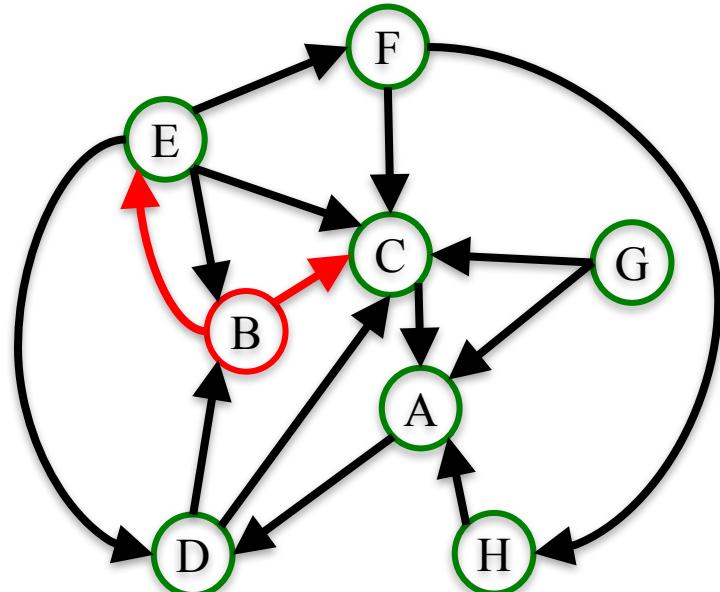
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	
C	1/3	1/15	12/15	
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



HITS Algorithm Example

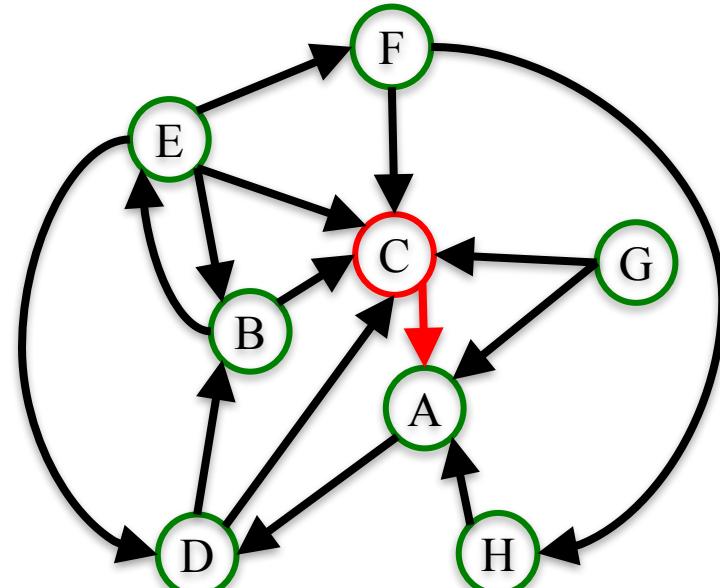
	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	2/5
C	1/3	1/15	12/15	
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



$$\frac{1}{3} + \frac{1}{15} = \frac{6}{15} = \frac{2}{5}$$

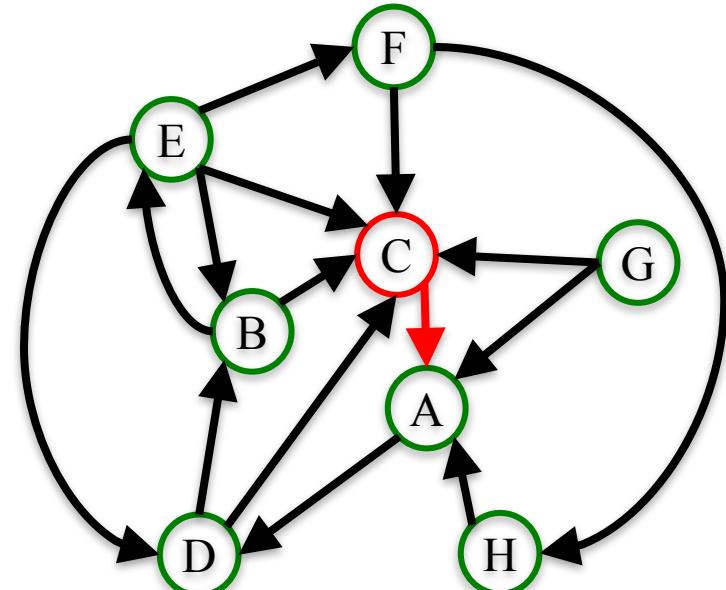
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	2/5
C	1/3	1/15	12/15	
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



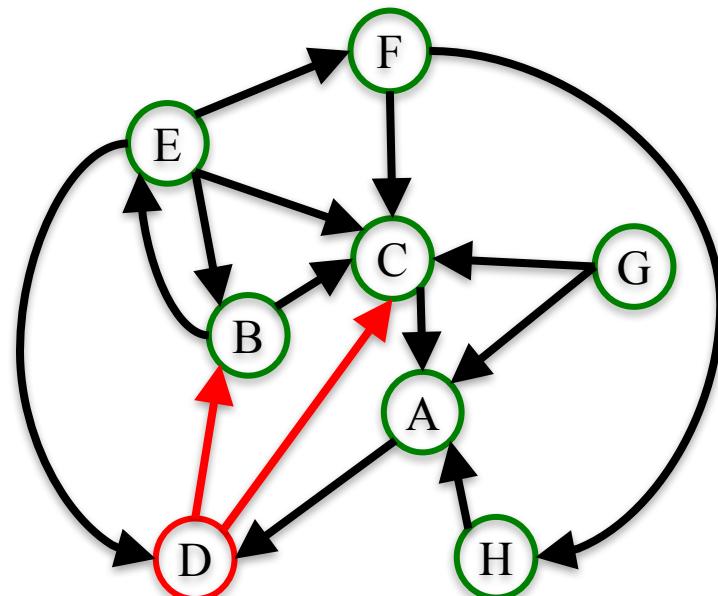
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	2/5
C	1/3	1/15	12/15	1/5
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



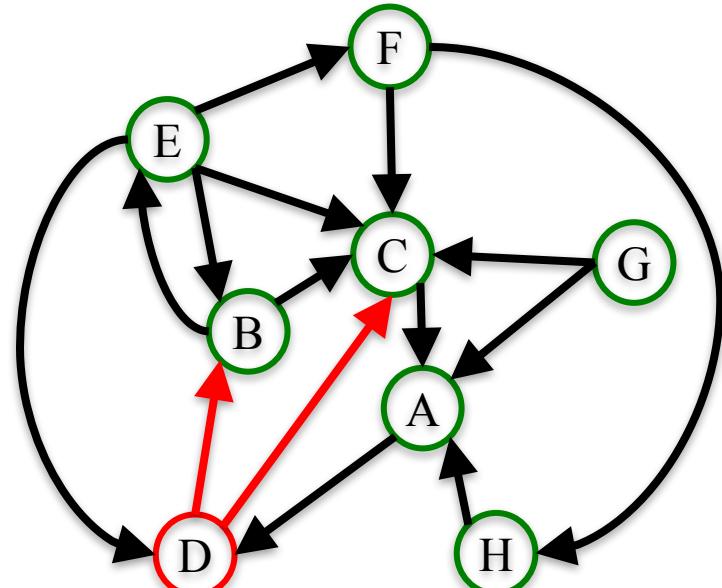
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	2/5
C	1/3	1/15	12/15	1/5
D	2/15	2/15	1/3	
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



HITS Algorithm Example

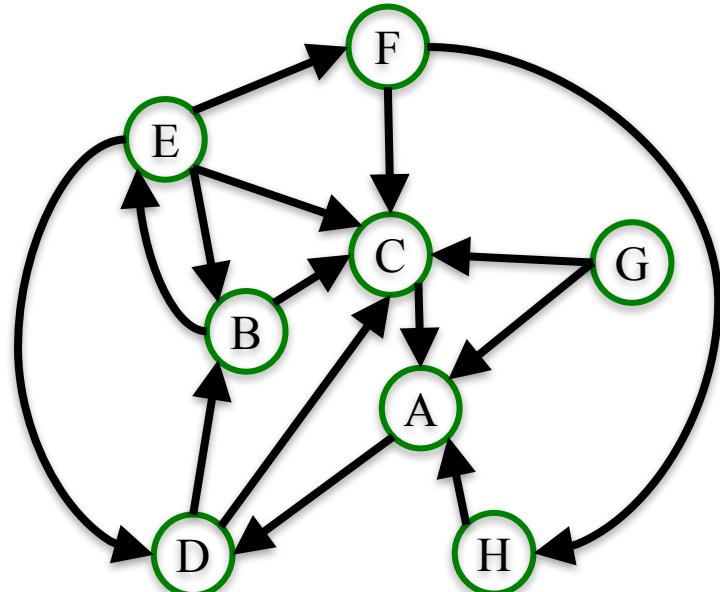
	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	2/5
C	1/3	1/15	12/15	1/5
D	2/15	2/15	1/3	7/15
E	1/15	4/15	2/15	
F	1/15	2/15	4/15	
G	0	2/15	0	
H	1/15	1/15	2/15	



$$2/15 + 1/3 = 7/15$$

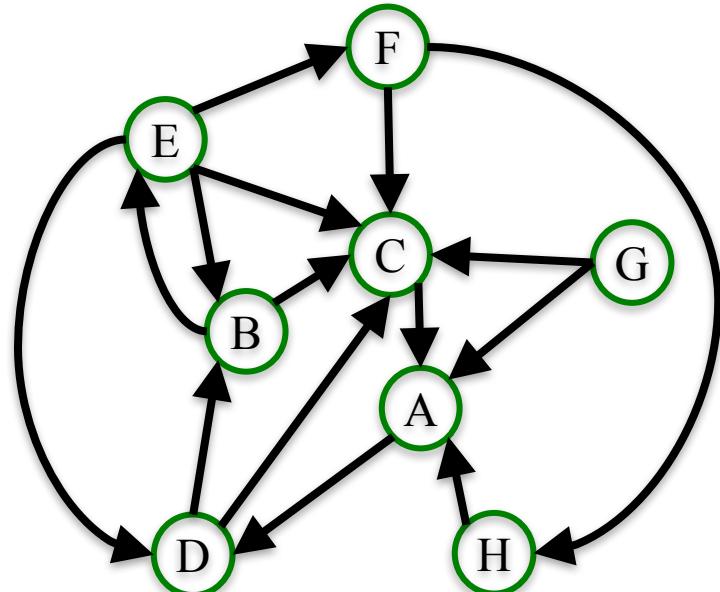
HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	2/5
C	1/3	1/15	12/15	1/5
D	2/15	2/15	1/3	7/15
E	1/15	4/15	2/15	2/3
F	1/15	2/15	4/15	2/5
G	0	2/15	0	8/15
H	1/15	1/15	2/15	1/5



HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/15	2/15
B	2/15	2/15	6/15	2/5
C	1/3	1/15	12/15	1/5
D	2/15	2/15	1/3	7/15
E	1/15	4/15	2/15	2/3
F	1/15	2/15	4/15	2/5
G	0	2/15	0	8/15
H	1/15	1/15	2/15	1/5

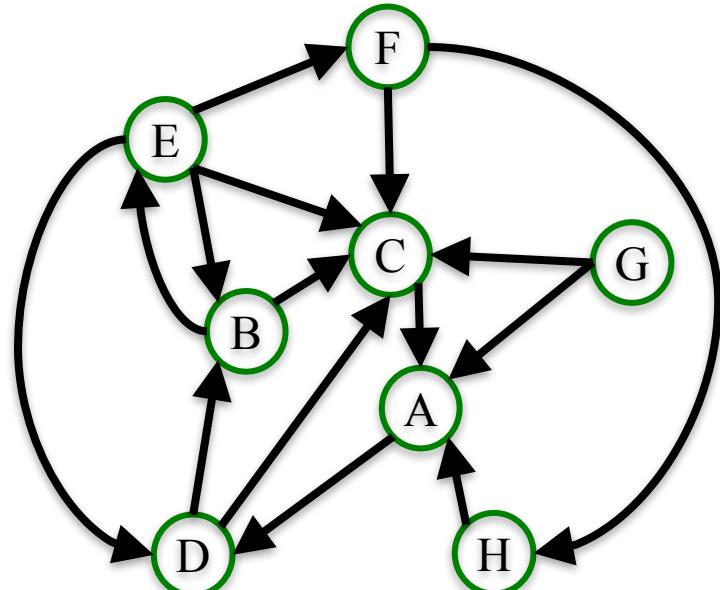


Normalize:

$$\sum_{i \in N} \text{auth}(i) = 35/15$$

HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/35	2/15
B	2/15	2/15	6/35	2/5
C	1/3	1/15	12/35	1/5
D	2/15	2/15	1/7	7/15
E	1/15	4/15	2/35	2/3
F	1/15	2/15	4/35	2/5
G	0	2/15	0	8/15
H	1/15	1/15	2/35	1/5

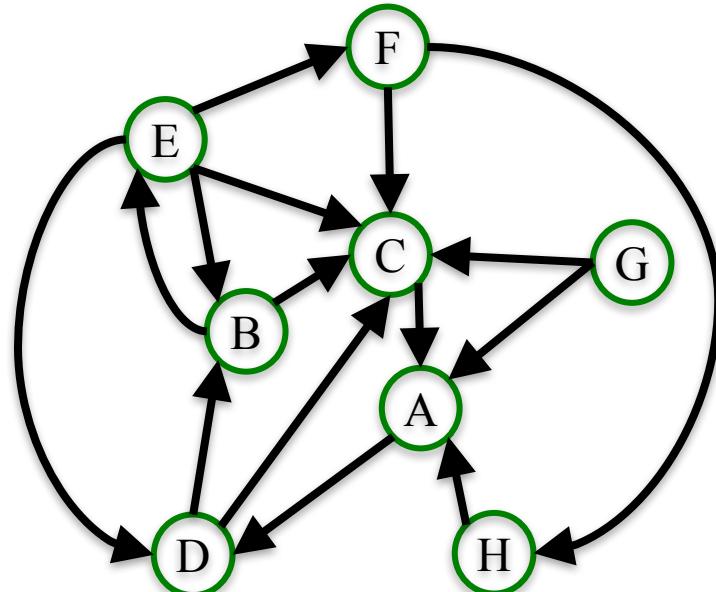


Normalize:

$$\sum_{i \in N} \text{auth}(i) = 35/15$$

HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/35	2/15
B	2/15	2/15	6/35	2/5
C	1/3	1/15	12/35	1/5
D	2/15	2/15	1/7	7/15
E	1/15	4/15	2/35	2/3
F	1/15	2/15	4/35	2/5
G	0	2/15	0	8/15
H	1/15	1/15	2/35	1/5

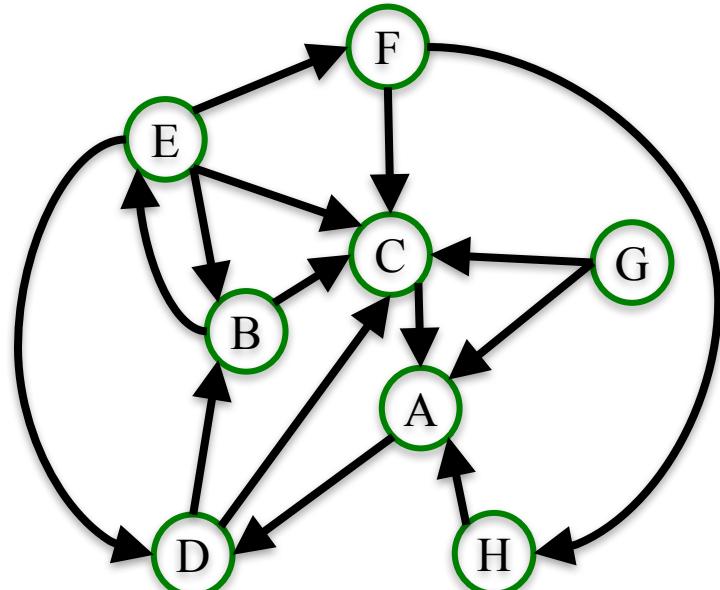


Normalize:

$$\sum_{i \in N} \text{hub}(i) = 45/15 = 3$$

HITS Algorithm Example

	Old Auth	Old Hub	New Auth	New Hub
A	1/5	1/15	4/35	2/45
B	2/15	2/15	6/35	2/15
C	1/3	1/15	12/35	1/15
D	2/15	2/15	1/7	7/45
E	1/15	4/15	2/35	2/9
F	1/15	2/15	4/35	2/15
G	0	2/15	0	8/45
H	1/15	1/15	2/35	1/15

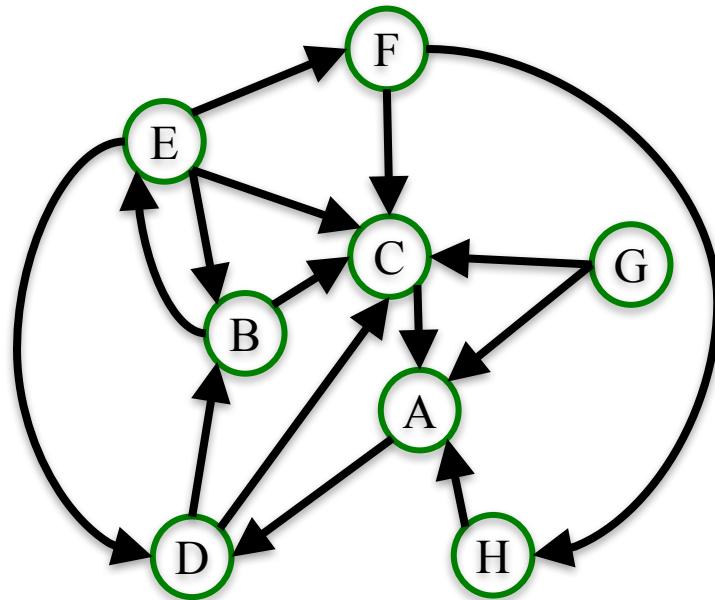


Normalize:

$$\sum_{i \in N} \text{hub}(i) = 45 / 15 = 3$$

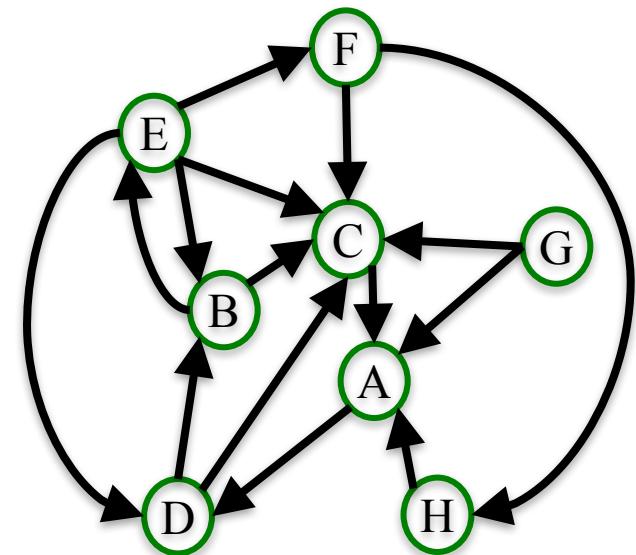
HITS Algorithm Convergence

What happens to the scores if we continue iterating the algorithm?



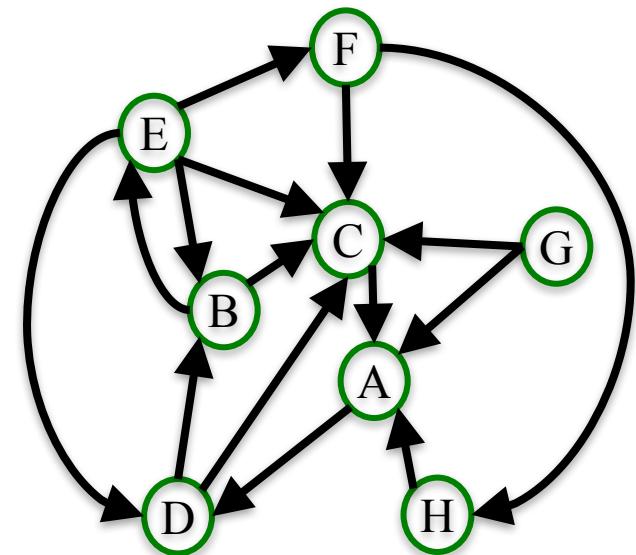
HITS Algorithm Convergence

	k	A	B	C	D	E	F	G	H
Auth	2	.11	.17	.34	.14	.06	.11	0	.06



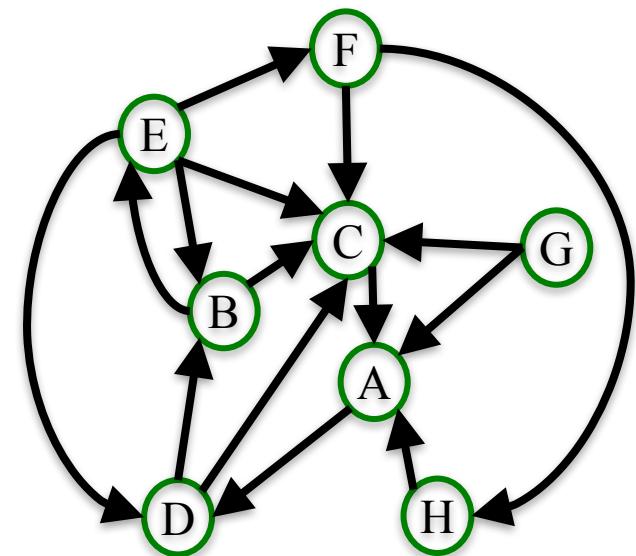
HITS Algorithm Convergence

	k	A	B	C	D	E	F	G	H
Auth	2	.11	.17	.34	.14	.06	.11	0	.06
	4	.10	.18	.36	.13	.06	.11	0	.06



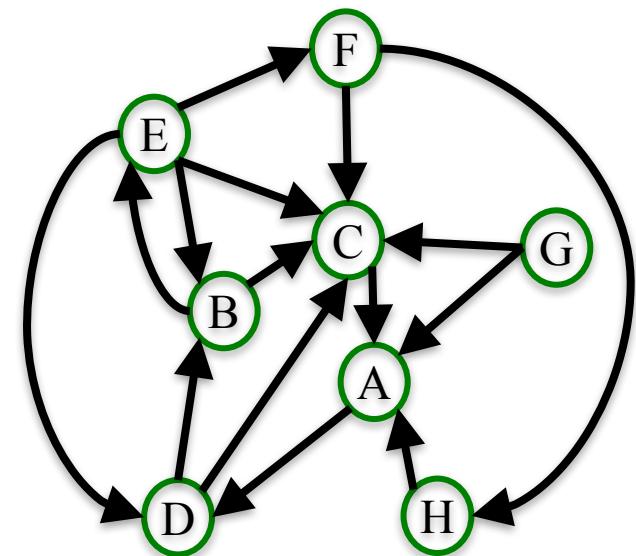
HITS Algorithm Convergence

	k	A	B	C	D	E	F	G	H
Auth	2	.11	.17	.34	.14	.06	.11	0	.06
	4	.10	.18	.36	.13	.06	.11	0	.06
	6	.09	.19	.37	.13	.06	.11	0	.06



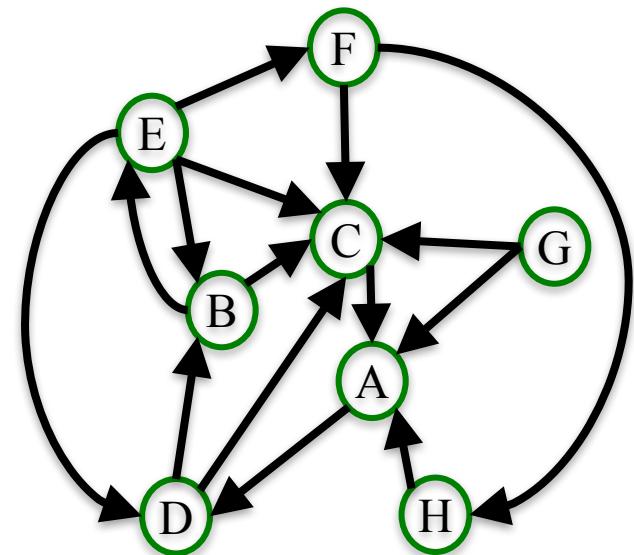
HITS Algorithm Convergence

	<i>k</i>	A	B	C	D	E	F	G	H
Auth	2	.11	.17	.34	.14	.06	.11	0	.06
	4	.10	.18	.36	.13	.06	.11	0	.06
	6	.09	.19	.37	.13	.06	.11	0	.06
Hub	2	.04	.13	.07	.16	.22	.13	.18	.07



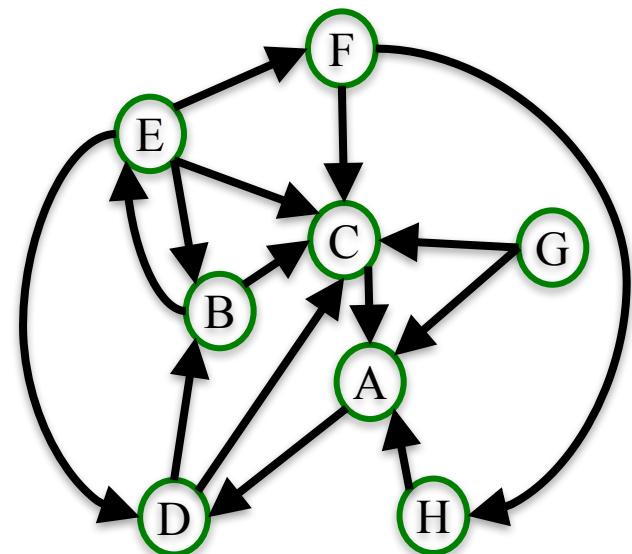
HITS Algorithm Convergence

	<i>k</i>	A	B	C	D	E	F	G	H
Auth	2	.11	.17	.34	.14	.06	.11	0	.06
	4	.10	.18	.36	.13	.06	.11	0	.06
	6	.09	.19	.37	.13	.06	.11	0	.06
Hub	2	.04	.13	.07	.16	.22	.13	.18	.07
	4	.04	.14	.05	.18	.25	.14	.17	.04



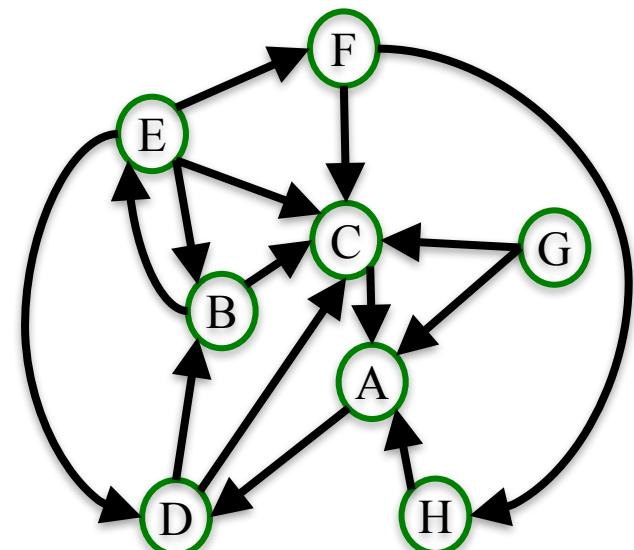
HITS Algorithm Convergence

	<i>k</i>	A	B	C	D	E	F	G	H
Auth	2	.11	.17	.34	.14	.06	.11	0	.06
	4	.10	.18	.36	.13	.06	.11	0	.06
	6	.09	.19	.37	.13	.06	.11	0	.06
Hub	2	.04	.13	.07	.16	.22	.13	.18	.07
	4	.04	.14	.05	.18	.25	.14	.17	.04
	6	.04	.14	.04	.18	.26	.14	.16	.04

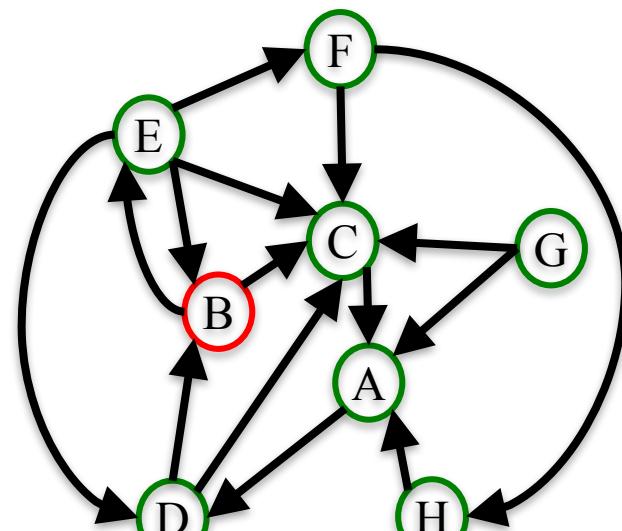
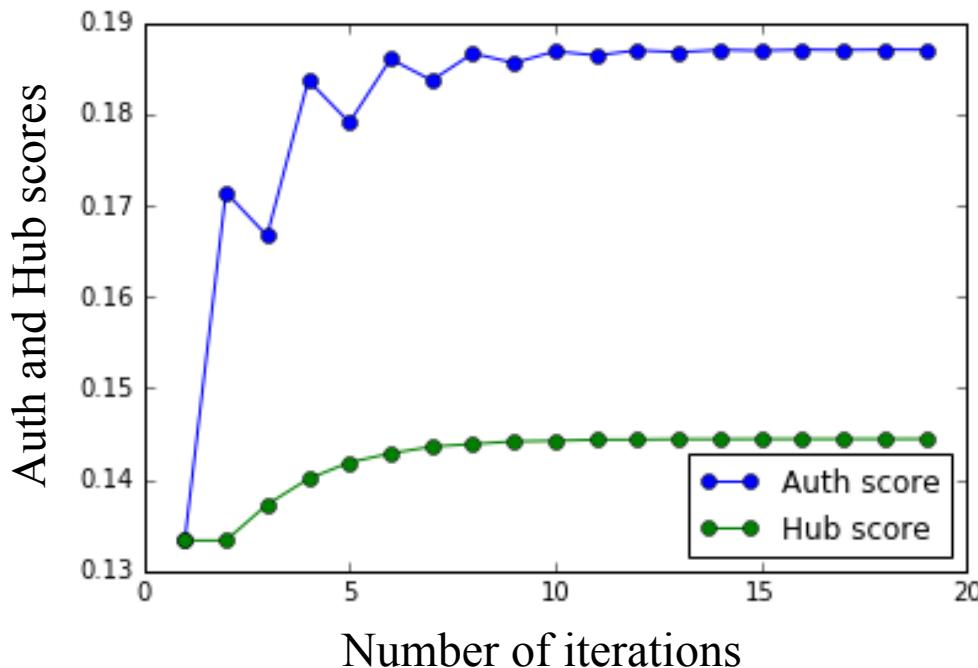


HITS Algorithm Convergence

	k	A	B	C	D	E	F	G	H
Auth	2	.11	.17	.34	.14	.06	.11	0	.06
	4	.10	.18	.36	.13	.06	.11	0	.06
	6	.09	.19	.37	.13	.06	.11	0	.06
Hub	2	.04	.13	.07	.16	.22	.13	.18	.07
	4	.04	.14	.05	.18	.25	.14	.17	.04
	6	.04	.14	.04	.18	.26	.14	.16	.04

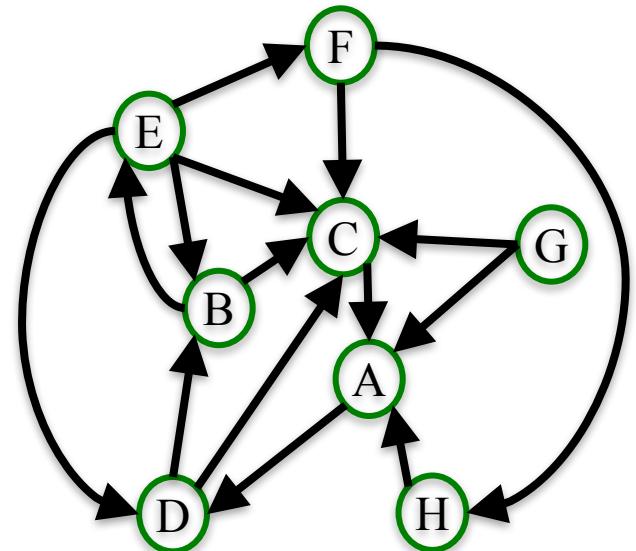


HITS Algorithm Convergence



HITS Algorithm Convergence

For most networks, as k gets larger, authority and hub scores converge to a unique value.

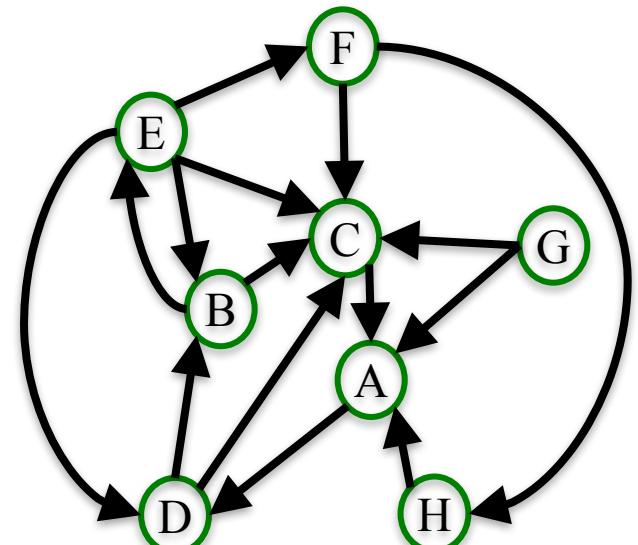


HITS Algorithm Convergence

For most networks, as k gets larger, authority and hub scores converge to a unique value.

As $k \rightarrow \infty$ the hub and authority scores approach:

	A	B	C	D	E	F	G	H
Auth	.08	.19	.40	.13	.06	.11	0	.06
Hub	.04	.14	.03	.19	.27	.14	.15	.03

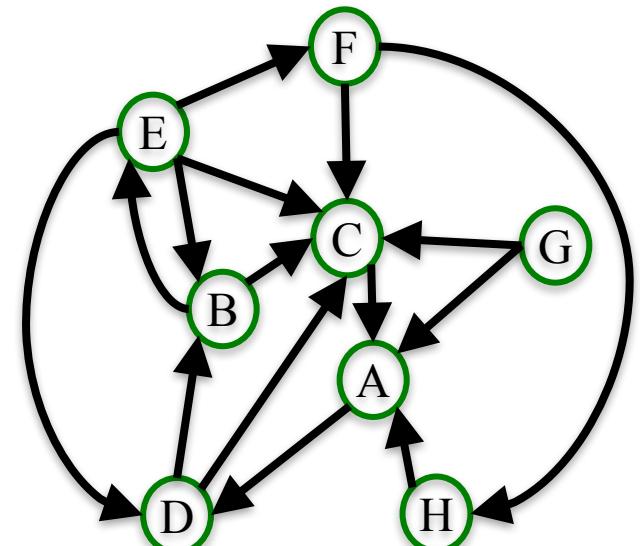


HITS Algorithm Convergence

For most networks, as k gets larger, authority and hub scores converge to a unique value.

As $k \rightarrow \infty$ the hub and authority scores approach:

	A	B	C	D	E	F	G	H
Auth	.08	.19	.40	.13	.06	.11	0	.06
Hub	.04	.14	.03	.19	.27	.14	.15	.03

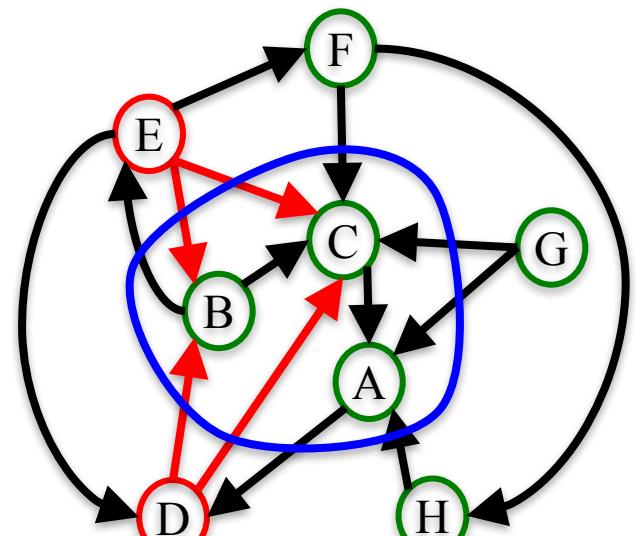


HITS Algorithm Convergence

For most networks, as k gets larger, authority and hub scores converge to a unique value.

As $k \rightarrow \infty$ the hub and authority scores approach:

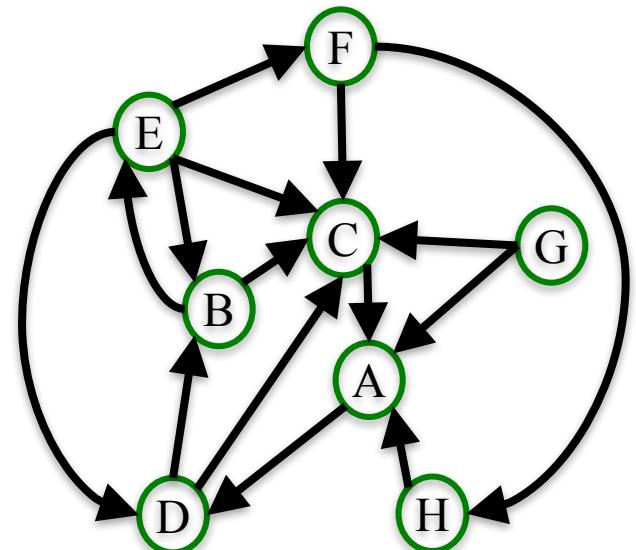
	A	B	C	D	E	F	G	H
Auth	.08	.19	.40	.13	.06	.11	0	.06
Hub	.04	.14	.03	.19	.27	.14	.15	.03



HITS Algorithm NetworkX

You can use NetworkX function `hits(G)` to compute the hub and authority scores of network G.

`hits(G)` outputs two dictionaries, keyed by node, with the hub and authority scores of the nodes.



Summary

- The HITS algorithm starts by constructing a *root* set of relevant web pages and expanding it to a *base* set.
- HITS then assigns an authority and hub score to each node in the network.
- Nodes that have incoming edges from *good hubs* are *good authorities*, and nodes that have outgoing edges to *good authorities* are *good hubs*.
- Authority and hub scores converge for most networks.
- You can use NetworkX function `hits(G)` to compute the hub and authority scores of network G

