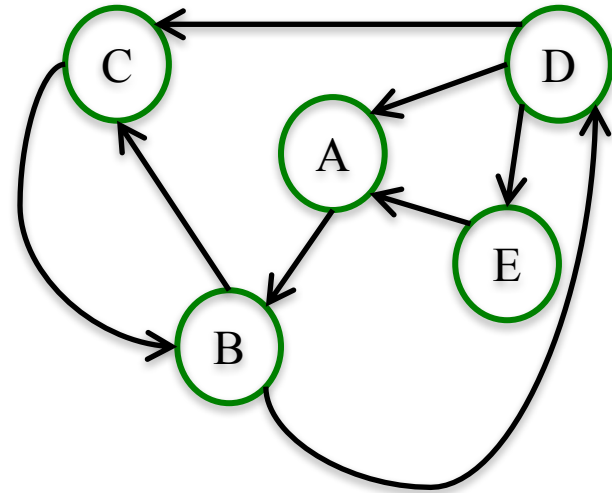


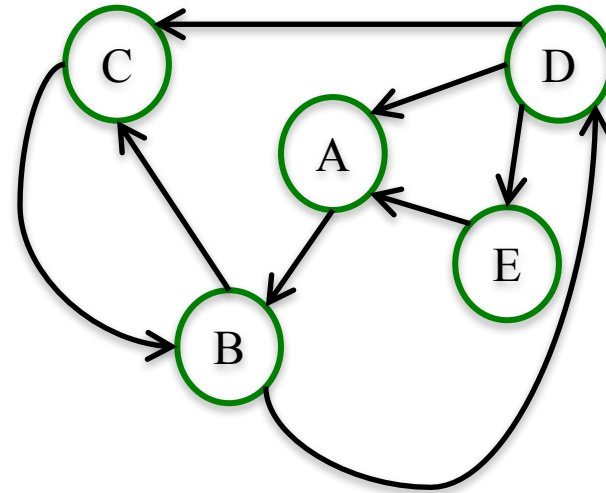
Interpreting PageRank

The PageRank of a node at step k is the probability that a **random walker** lands on the node after taking k steps.



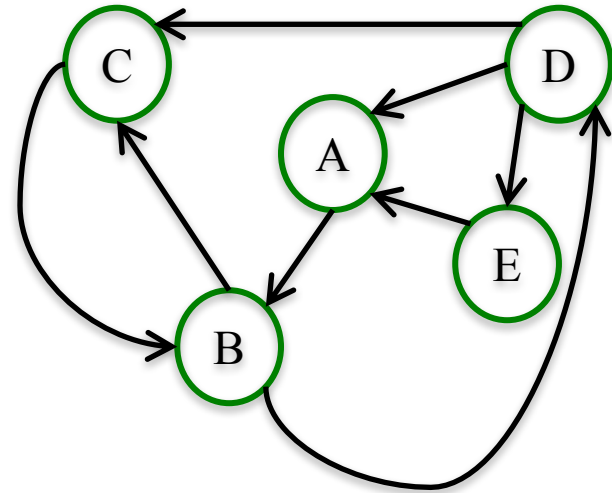
Interpreting PageRank

Random walk of k steps: Start on a random node.



Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node.

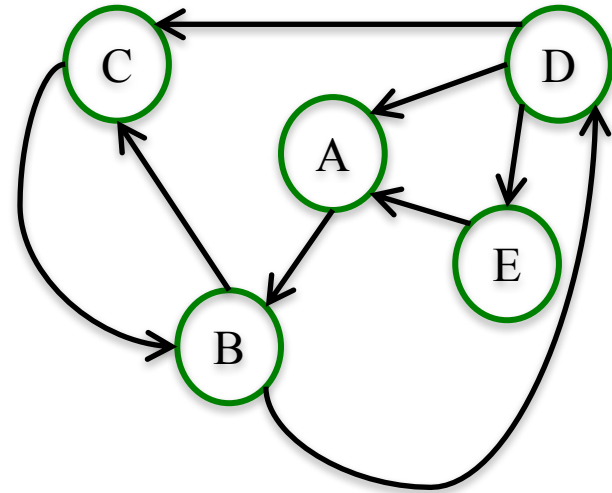


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 1: Choose a random node.

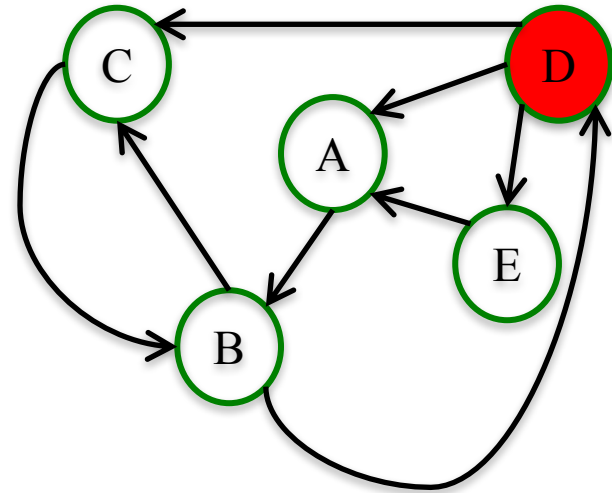


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 1: Choose a random node.

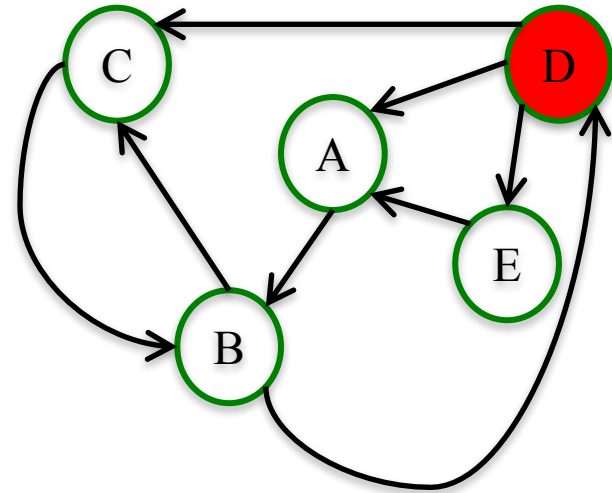


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 1: Choose a random outgoing edge.

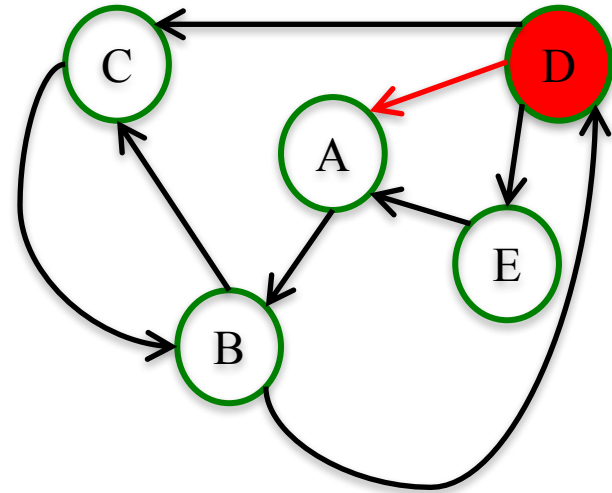


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 1: Choose a random outgoing edge.

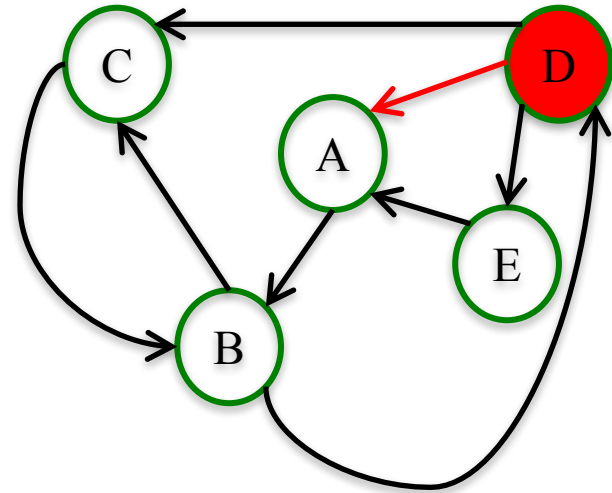


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 1: Follow the edge to the next node.

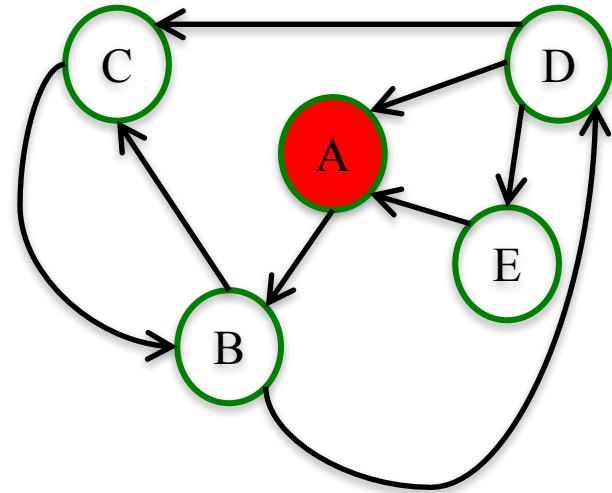


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 1: Follow the edge to the next node.

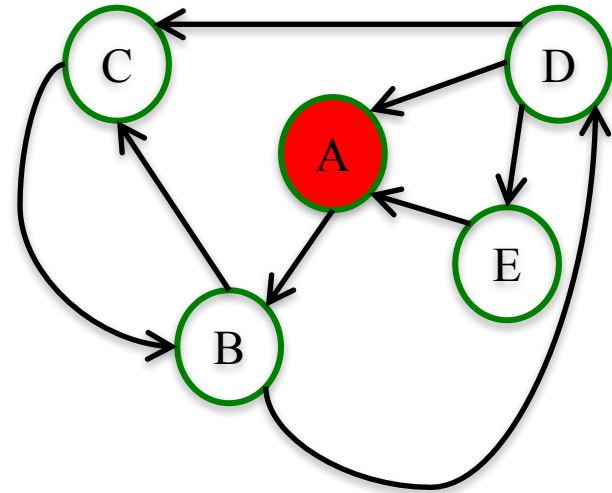


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 2: Choose a random outgoing edge and follow it.

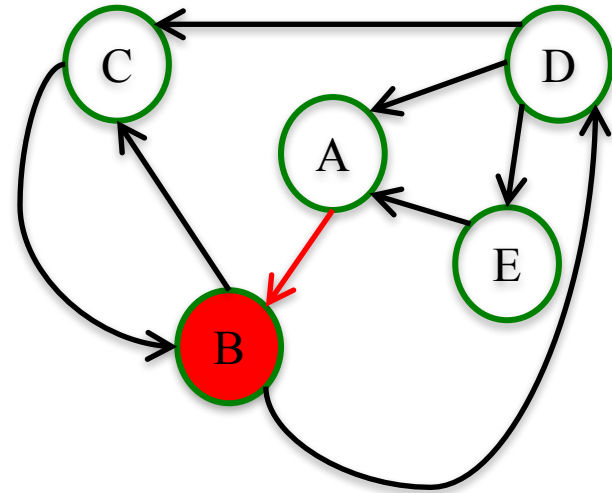


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 2: Choose a random outgoing edge and follow it.

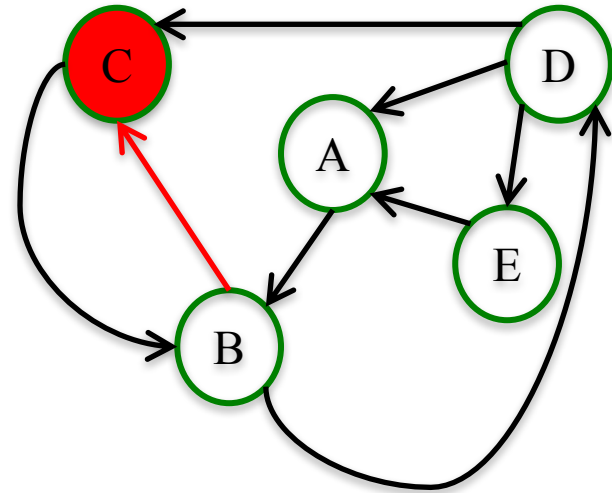


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 3: Choose a random outgoing edge and follow it.

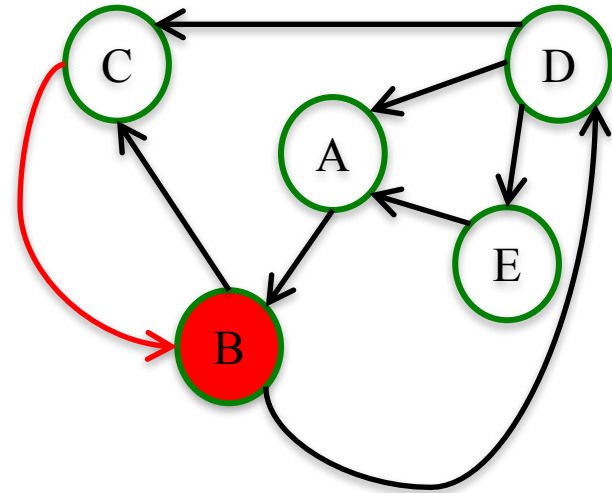


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

Step 4: Choose a random outgoing edge and follow it.

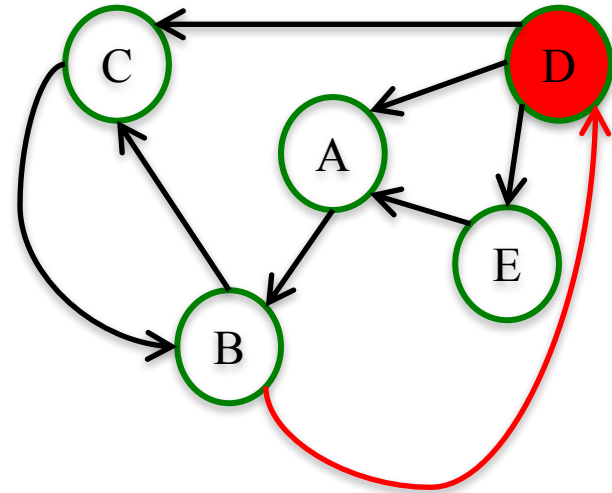


Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

For example, a random walk of 5 steps on this graph looks like this:

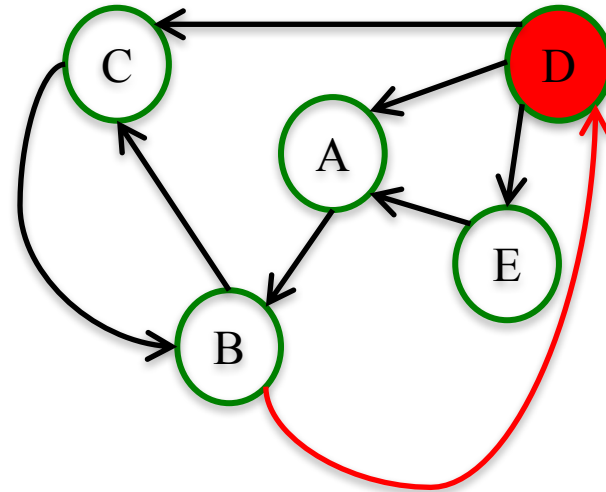
Step 5: Choose a random outgoing edge and follow it.



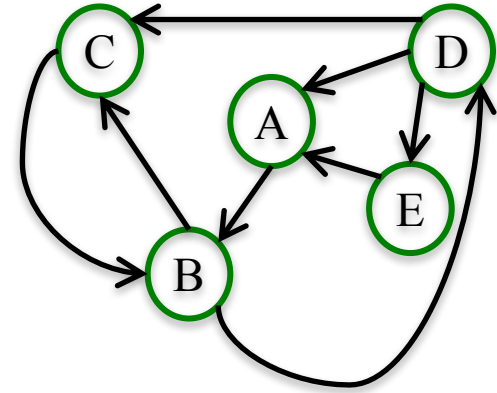
Interpreting PageRank

Random walk of k steps: Start on a random node. Then choose an outgoing edge at random and follow it to the next node. Repeat k times.

	Page Rank				
	A	B	C	D	E
$k=\infty$.12	.38	.25	.19	.06



PageRank Problem



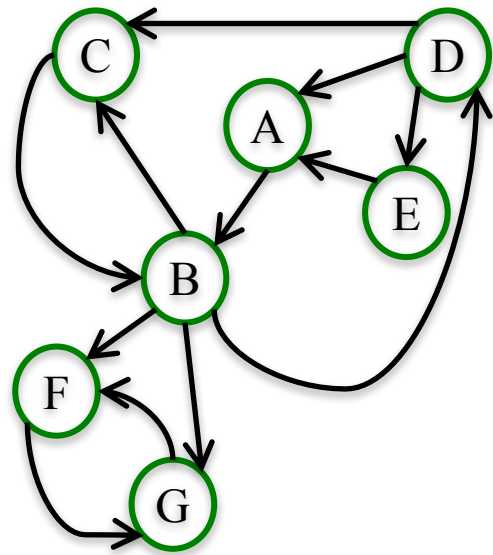
PageRank Problem

What's the PageRank of the nodes in this network?
[Hint: think about the random walk interpretation]

For a large enough k : F and G each have PageRank of $\frac{1}{2}$ and all the other nodes have PageRank 0.

Why? Imagine a random walk on this network.
Whenever the walk lands on F or G, it is “stuck” on F and G.

This seems problematic!



PageRank Problem

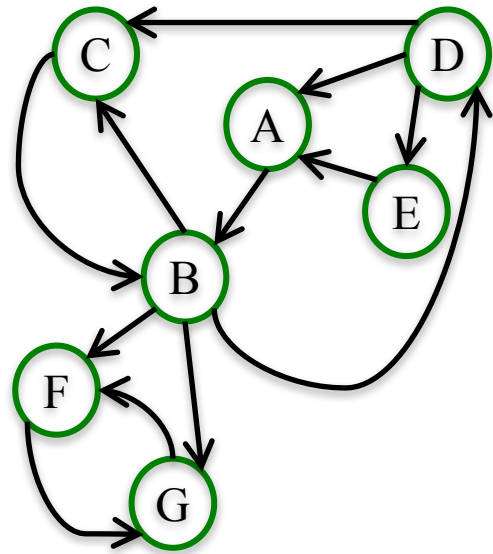
To fix this, we introduce a “damping parameter” α .

Random walk of k steps with damping parameter α : Start on a random node. Then:

- **With probability α :** choose an outgoing edge at random and follow it to the next node.
- **With probability $1 - \alpha$:** choose a node at random and go to it.

Repeat k times.

The random walk is no longer “stuck” on nodes F and G.

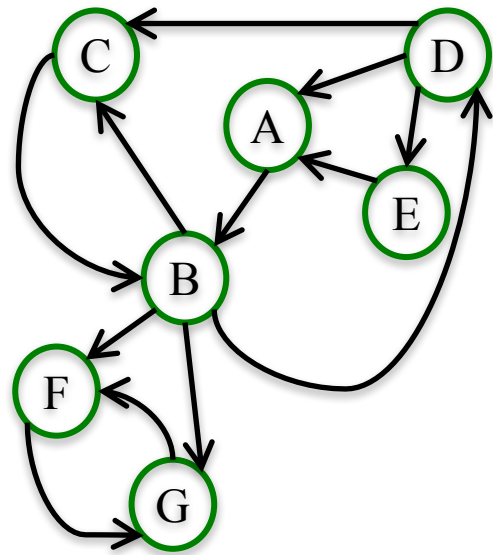


Scaled PageRank

The **Scaled PageRank** of k steps and damping factor α of a node n is the probability that a random walk with damping factor α lands on a n after k steps.

For most networks, as k gets larger, Scaled PageRank converges to a unique value, which depends on α .

In practice, we use a parameter of α between 0.8 and 0.9.



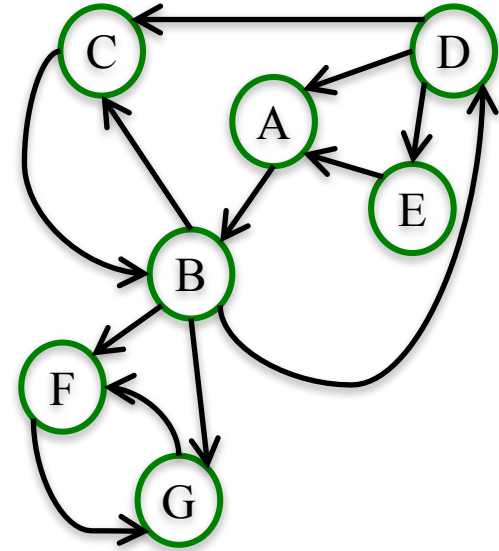
Scaled PageRank

Scaled PageRank ($\alpha = .8, k$ large)						
A	B	C	D	E	F	G
.08	.17	.1	.08	.05	.27	.25

F and G still have high PageRank, but not all the PageRank.

Damping factor works better in very large networks like the Web or large social networks.

You can use NetworkX function `pagerank(G, alpha=0.8)` to compute Scaled PageRank of network G with damping parameter alpha.



Summary

- The Basic PageRank of a node can be interpreted as the probability that a random walk lands on the node after k random steps.
- Basic PageRank has the problem that, in some networks, a few nodes can “suck up” all the PageRank from the network.
- To fix this problem, Scaled PageRank introduces a parameter α , such that the random walker chooses a random node to jump to with probability $1 - \alpha$.
- Typically we use α between 0.8 and 0.9
- NetworkX function `pagerank(G, alpha=0.8)` computes Scaled PageRank of network G with damping parameter $\alpha=0.8$.

