Graph Ranking & The Cost of Sybil Defense

Gwen Farach-Colton
Rutgers University
USA

Martín Farach-Colton
Rutgers University
USA

Leslie Ann Goldberg
Oxford University
UK

Hanna Komlós
Rutgers University
USA

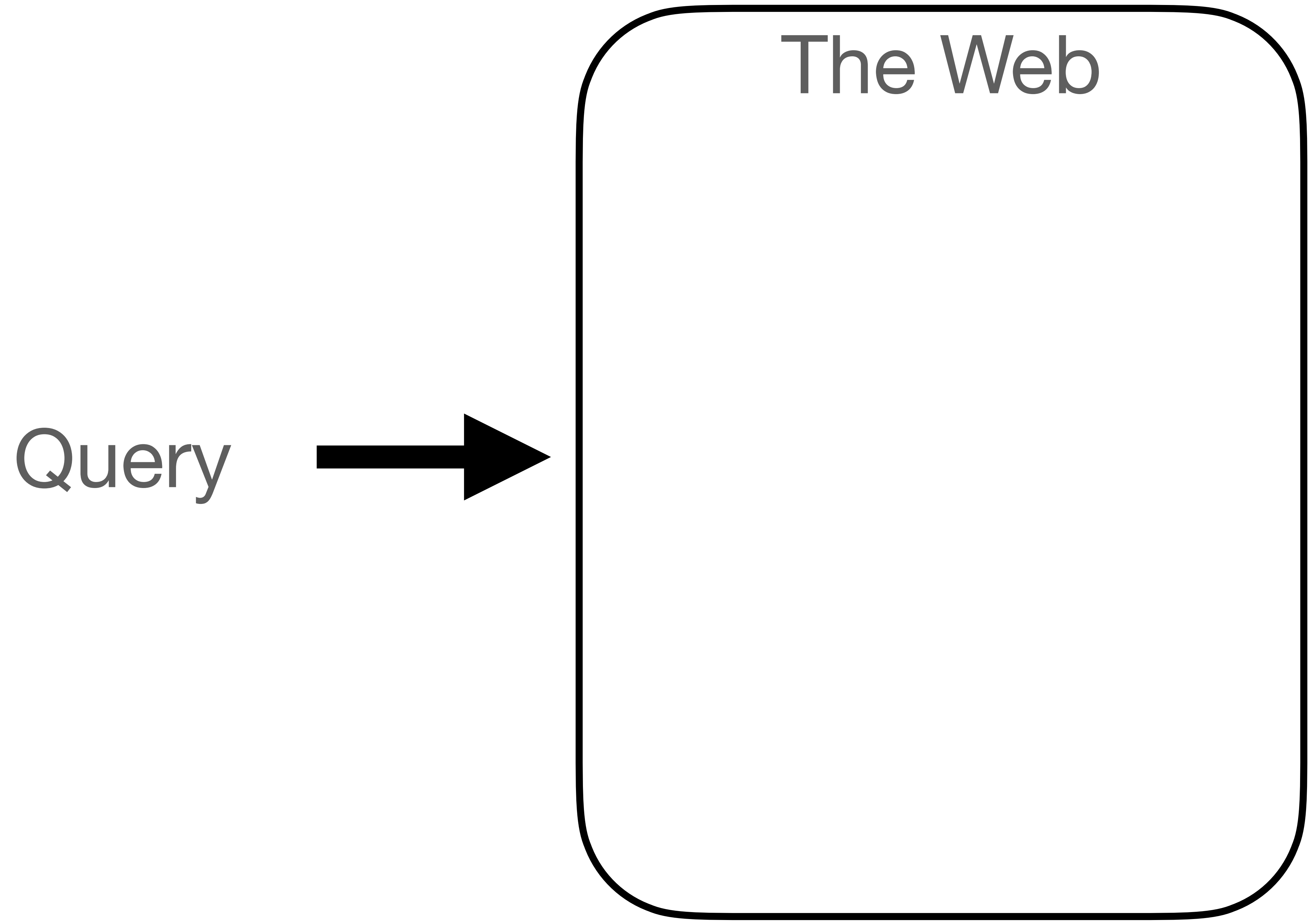
John Lapinskas
Bristol University
UK

Reut Levi
Reichman Univeristy
Israel

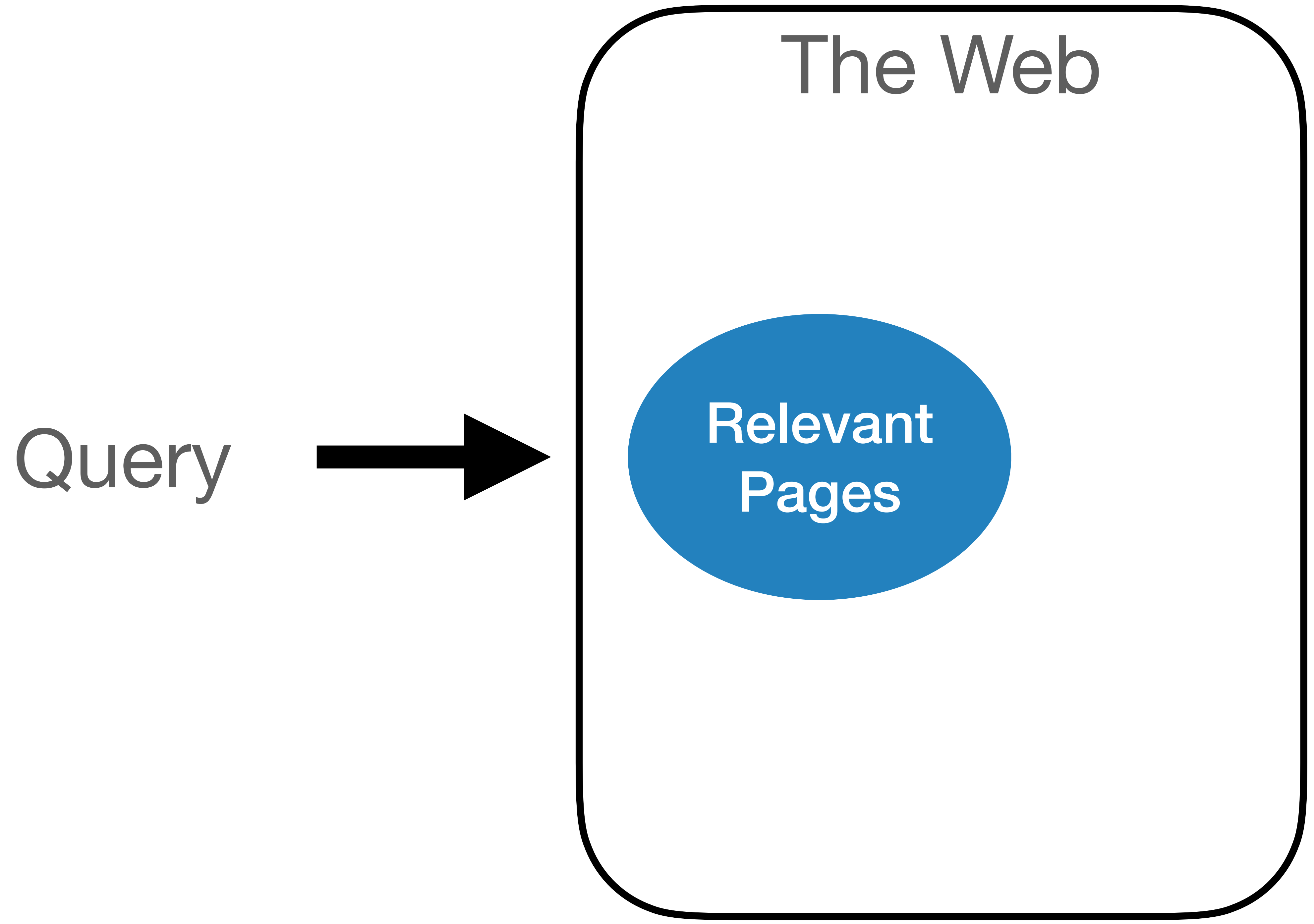
Moti Medina
Bar-Ilan University
Israel

Miguel Mosteiro
Pace University
USA

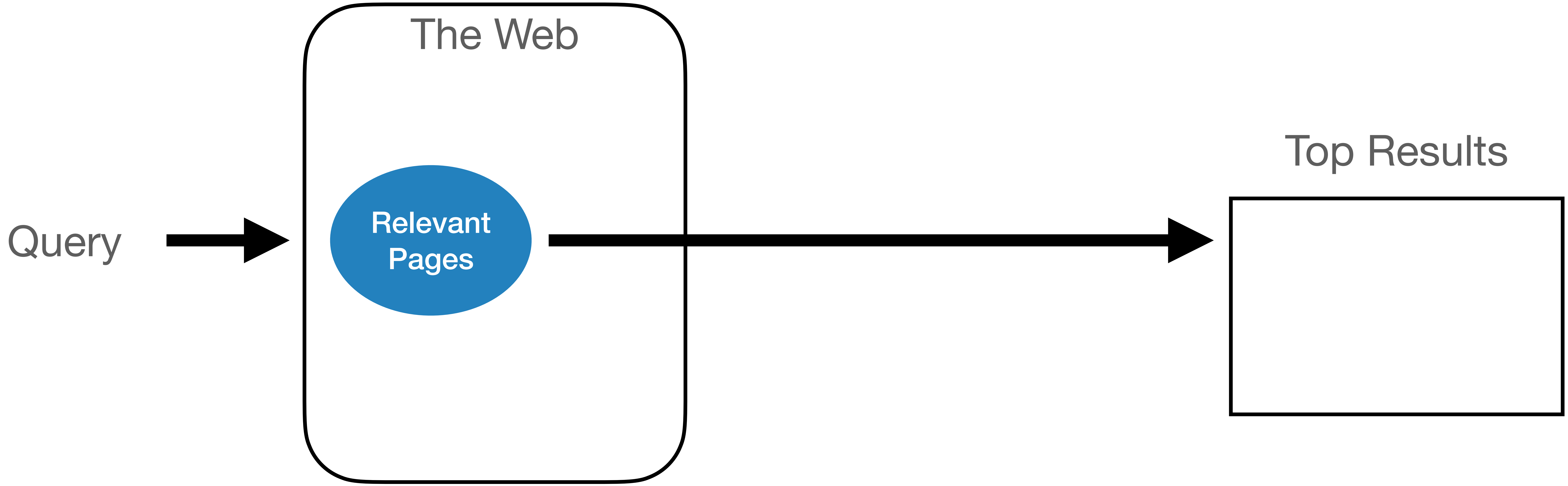
Schematic of how search engines work:



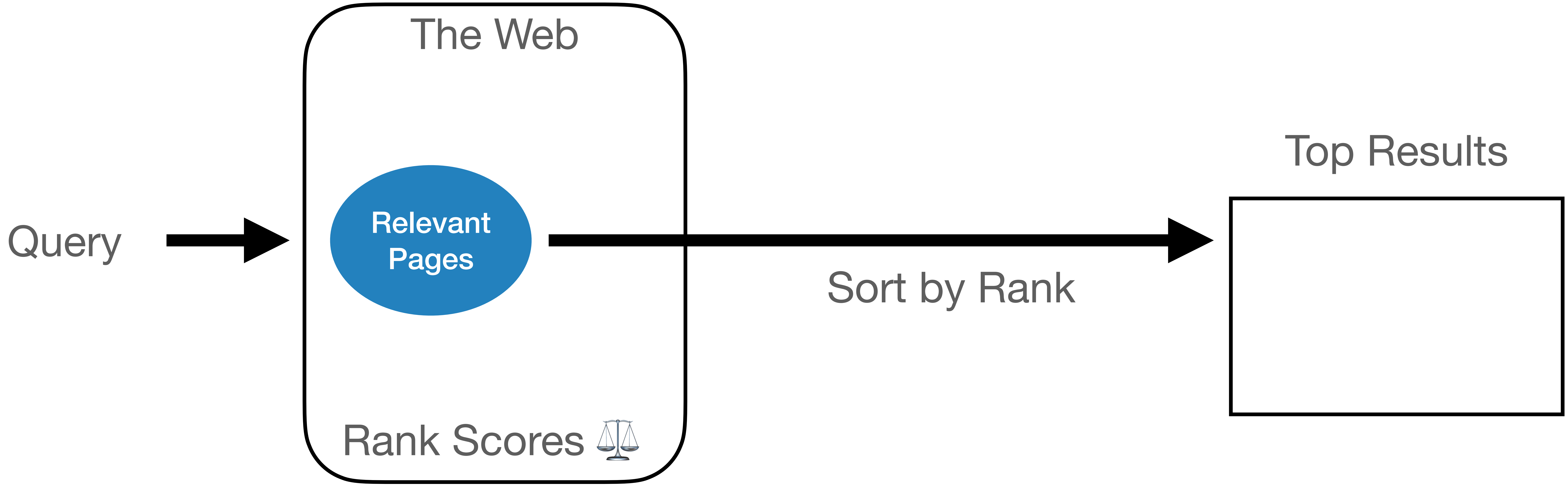
Schematic of how search engines work:



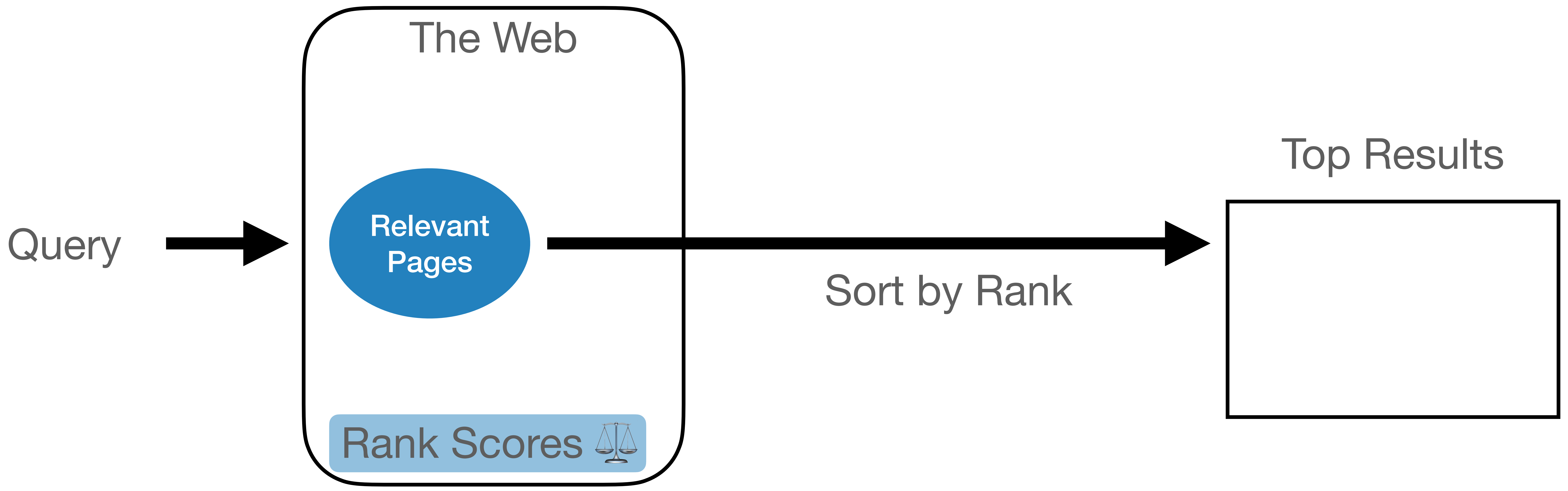
Schematic of how search engines work:



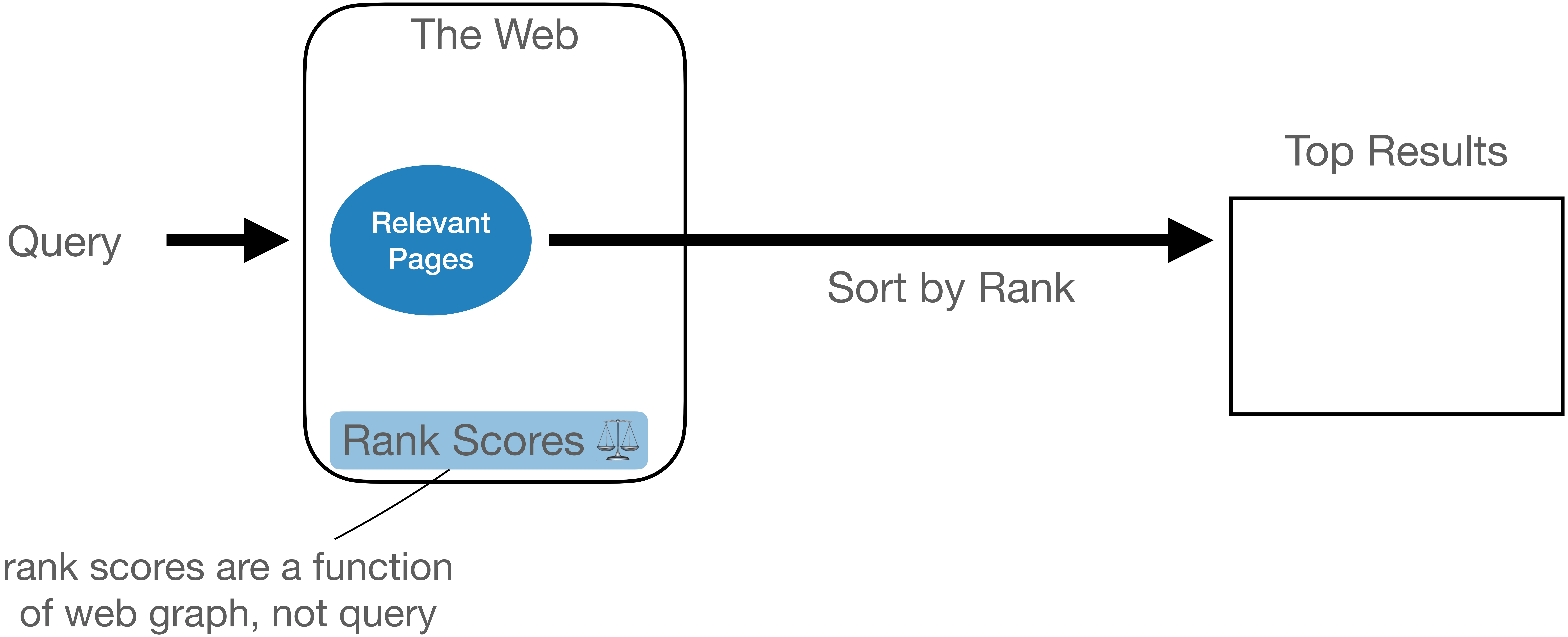
Schematic of how search engines work:



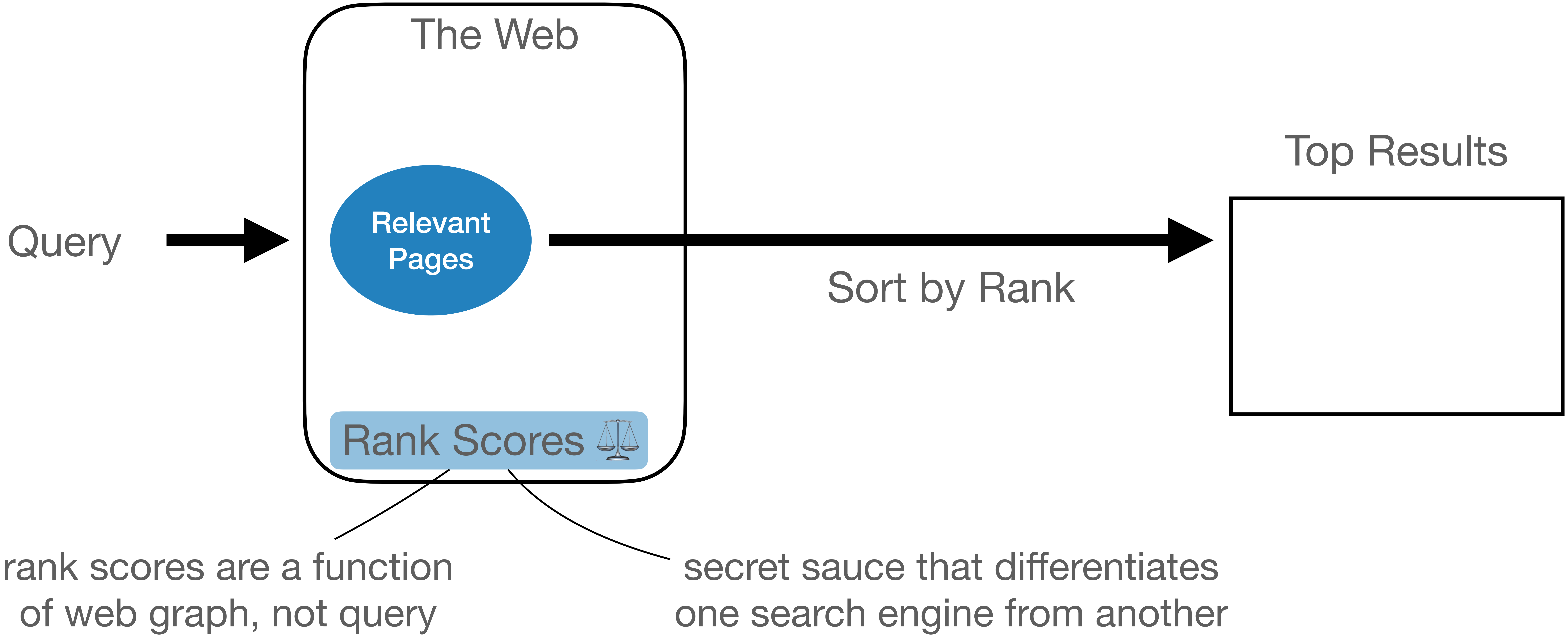
Schematic of how search engines work:



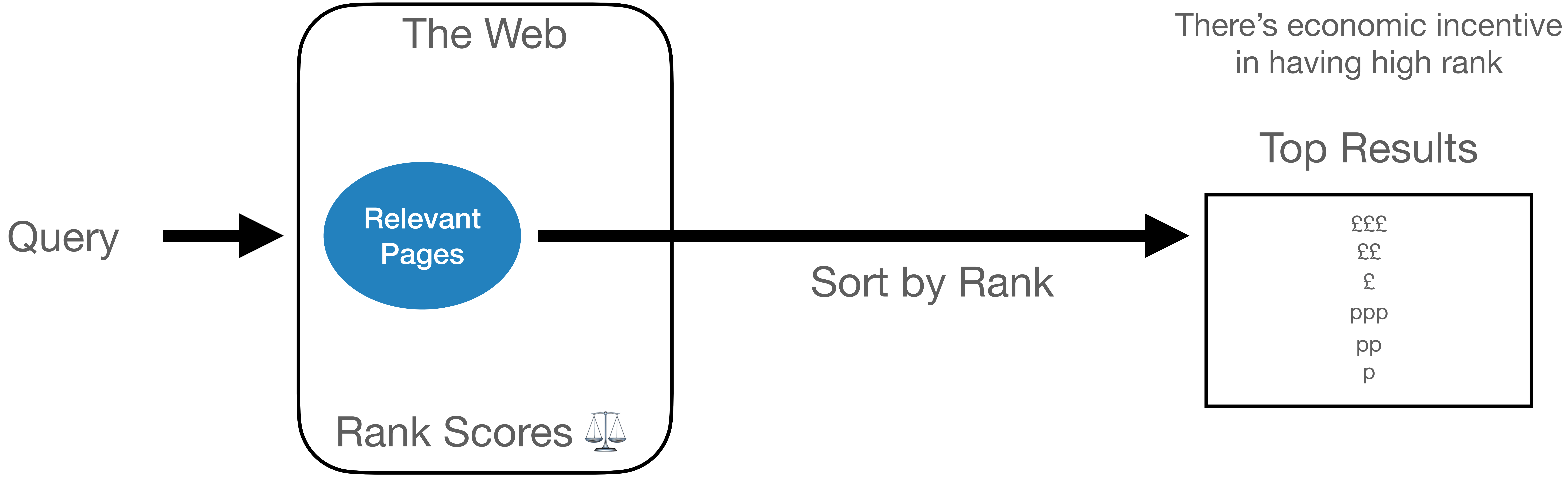
Schematic of how search engines work:



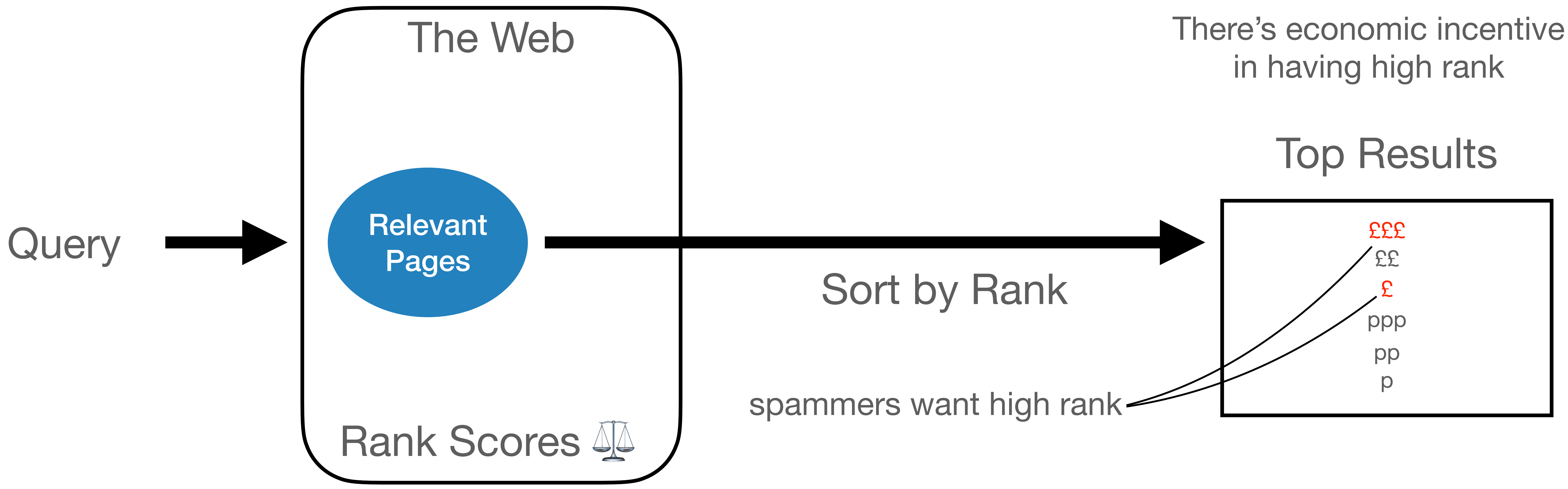
Schematic of how search engines work:



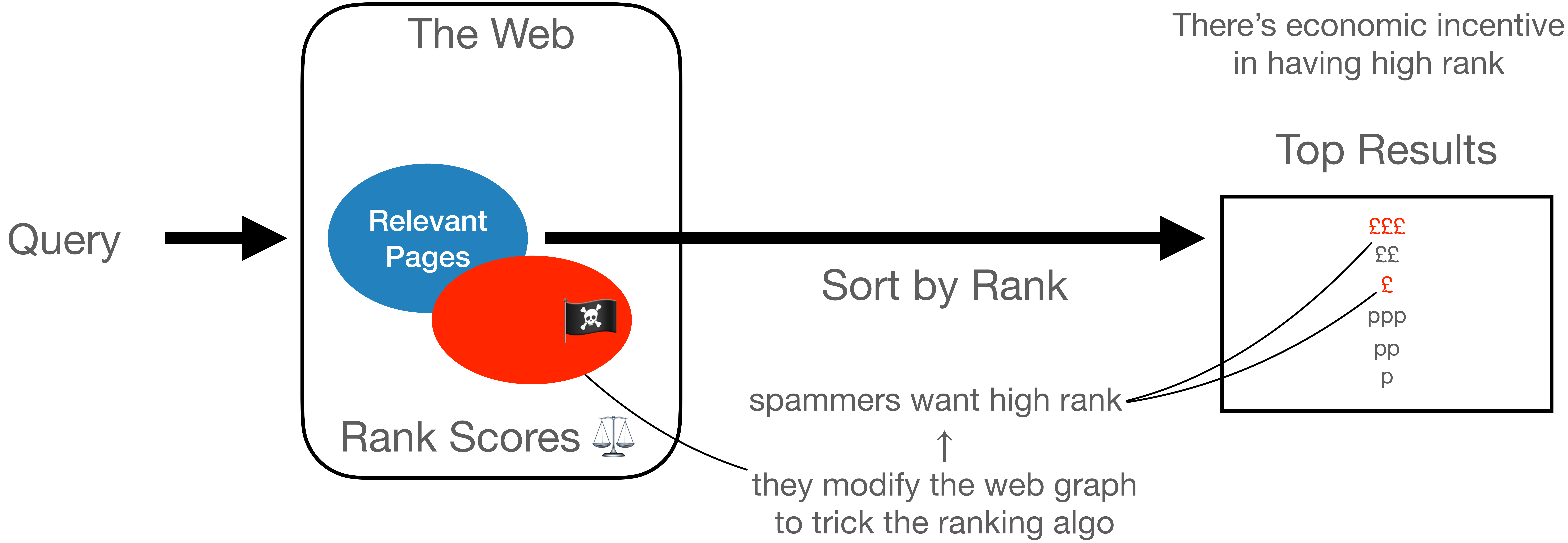
Schematic of how search engines work:



Schematic of how search engines work:



Schematic of how search engines work:



Our contribution

Formalize game between web ranker ⚖️ and spammer 🏴‍☠️

Define two critical criteria:

- **spam resistance**: how well the ranker can resist spamming
- **distortion**: a quality constraint on the ranker

Show that there is a ranking function with high resistance & low distortion!

- This is the version of PageRank actually used by Google

Lots of ancillary results about the algebra of PageRank

Formalize game between web ranker ⚖️ and spammer 🏴‍☠️

Define two critical criteria:

- **spam resistance**: how well the ranker can resist spamming
- **distortion**: a quality constraint on the ranker

Show that there is a ranking function with high resistance & low distortion!

- This is the version of PageRank actually used by Google

Lots of ancillary results about the algebra of PageRank

Our contribution

Formalize game between web ranker ⚖️ and spammer 🏴‍☠️

Define two critical criteria:

- **spam resistance**: how well the ranker can resist spamming
- **distortion**: a quality constraint on the ranker

Show that there is a ranking function with high resistance & low distortion!

- This is the version of PageRank actually used by Google

Lots of ancillary results about the algebra of PageRank

Our contribution

Formalize game between web ranker ⚖️ and spammer 🏴‍☠️

Define two critical criteria:

- **spam resistance**: how well the ranker can resist spamming
- **distortion**: a quality constraint on the ranker

Show that there is a ranking function with high resistance & low distortion!

- This is the version of PageRank actually used by Google

Lots of ancillary results about the algebra of PageRank

Running Example: PageRank

Where did PageRank come from?

Brin & Page’s insight:

- Good pages are pointed to by other good pages

How do you make that usable?

Where did PageRank come from?

Brin & Page's insight:

- Good pages are pointed to by other good pages

How do you make that usable?

Compute the Stationary Distribution of a Random Walk on the Web Graph

Where did PageRank come from?

Brin & Page's insight:

- Good pages are pointed to by other good pages

How do you make that usable?

Compute the Stationary Distribution of a Random Walk on the Web Graph

What's the relevance?

- We'll see that pages with high stationary distribution are pointed to by other pages with high stationary distribution

The Stationary Distribution of a Random Walk

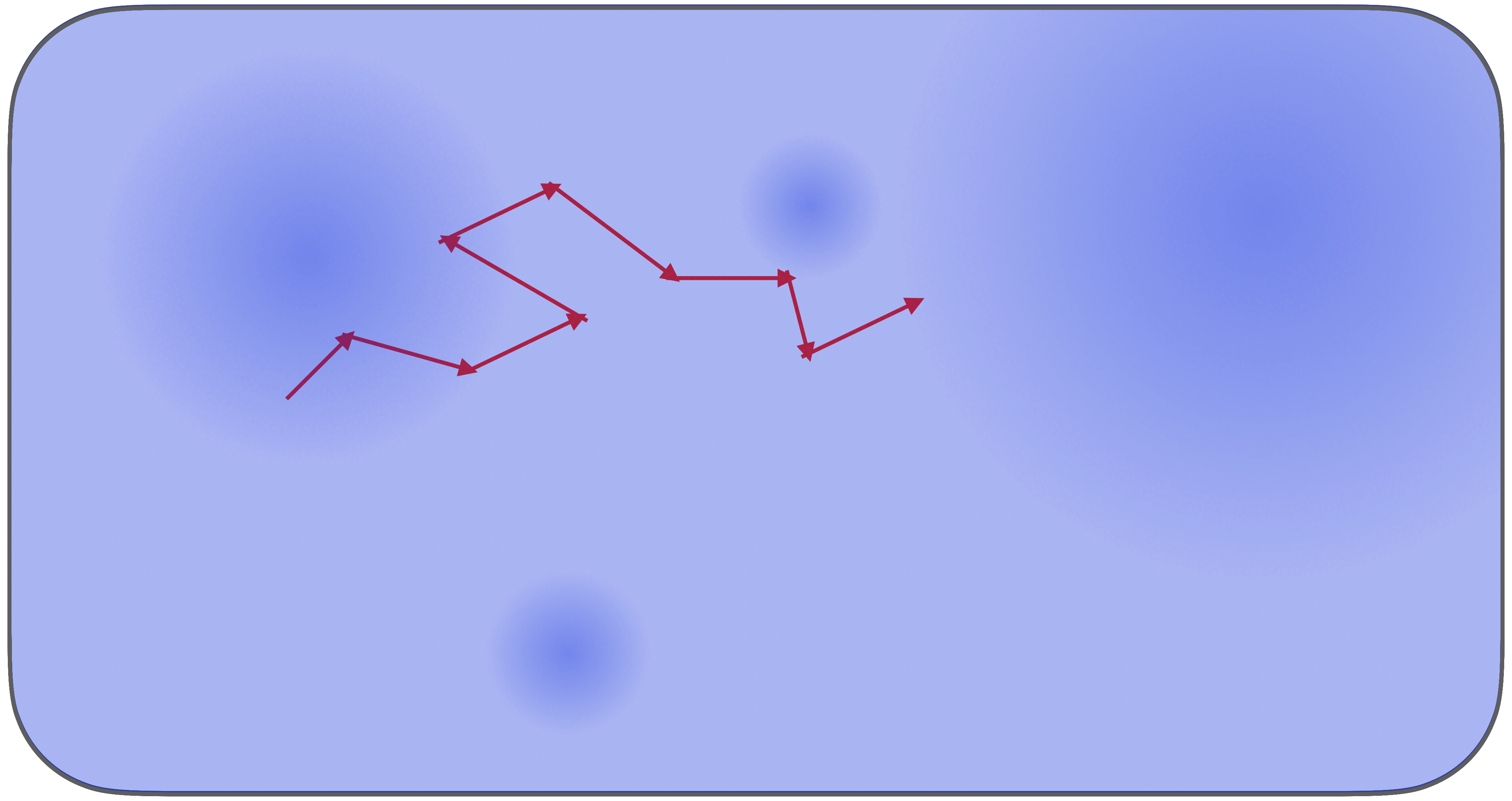
Definition:

Graph

1. Start at an arbitrary node
2. Take a random out edge
3. Go forever

Definition:

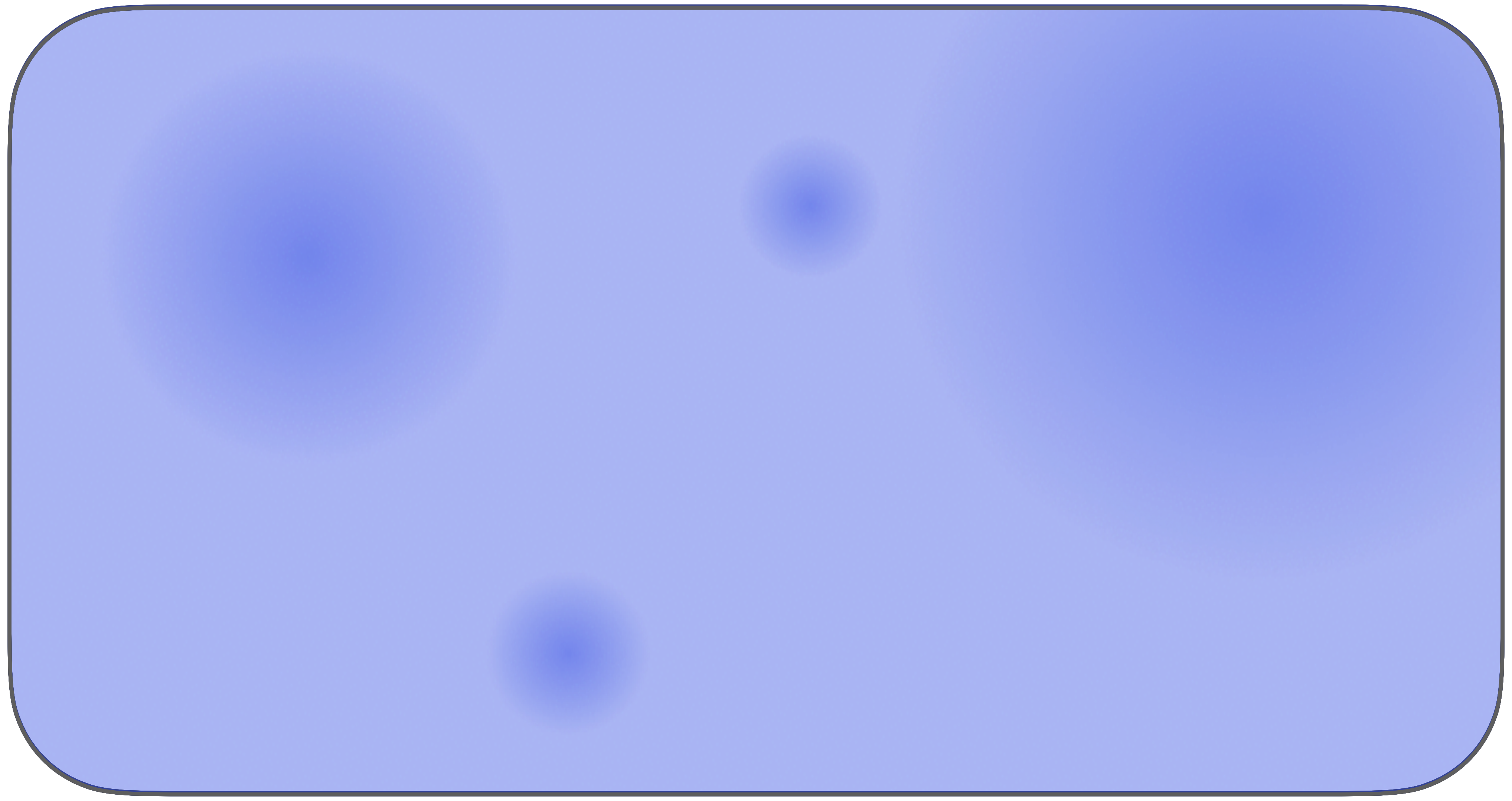
Graph



1. Start at an arbitrary node
2. Take a random out edge
3. Go forever

Definition:

Graph



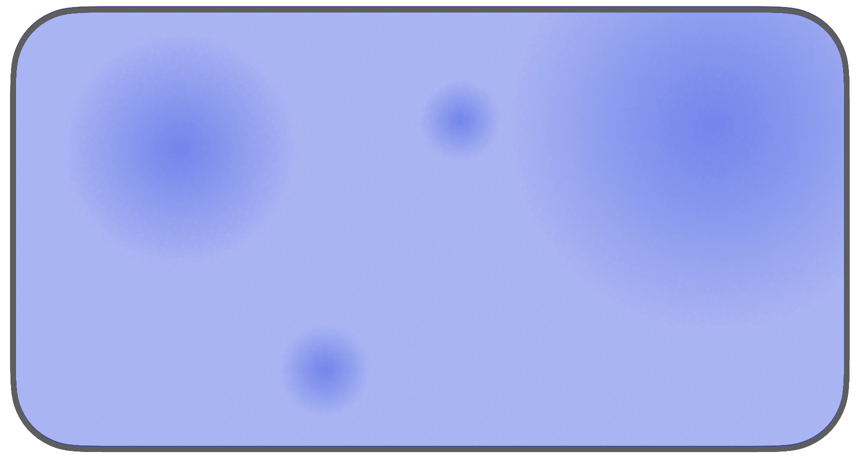
1. Start at an arbitrary node
2. Take a random out edge
3. Go forever

Fraction of time at each node = stationary distribution

Stationary Distribution: Pros and Cons

Pros:

- Reward having in-edges from high-quality pages
- Easy to compute



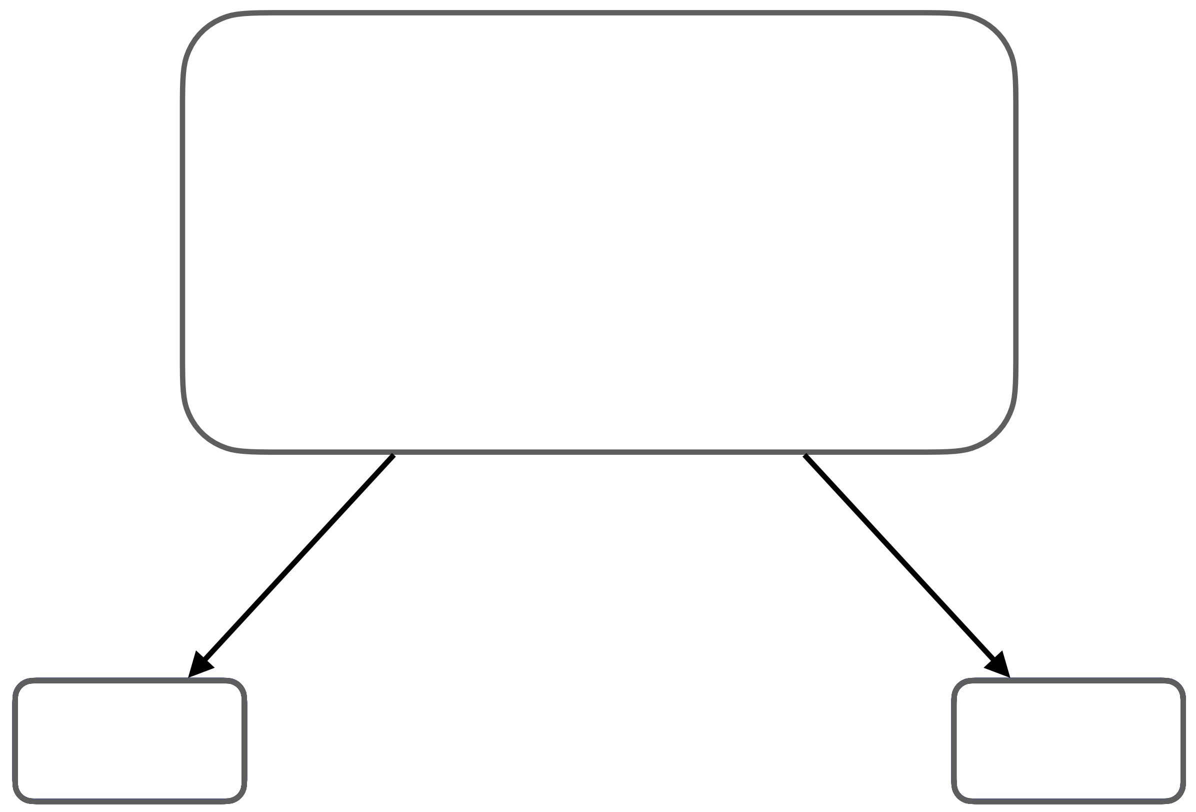
Stationary Distribution: Pros and Cons

Pros:

- Reward having in-edges from high-quality pages
- Easy to compute

Cons:

- Not defined on all graphs (only on ergodic graphs)
- Not defined for the web graph!



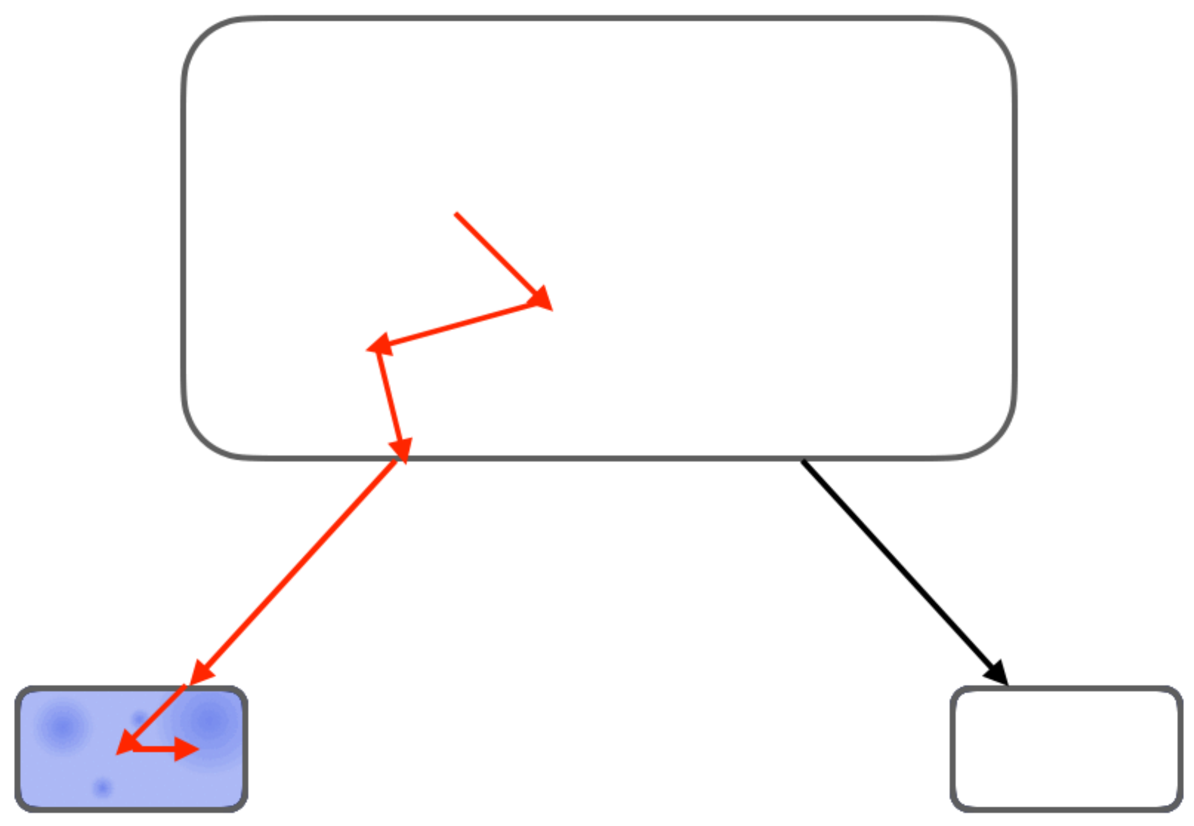
Stationary Distribution: Pros and Cons

Pros:

- Reward having in-edges from high-quality pages
- Easy to compute

Cons:

- Not defined on all graphs (only on ergodic graphs)
- Not defined for the web graph!



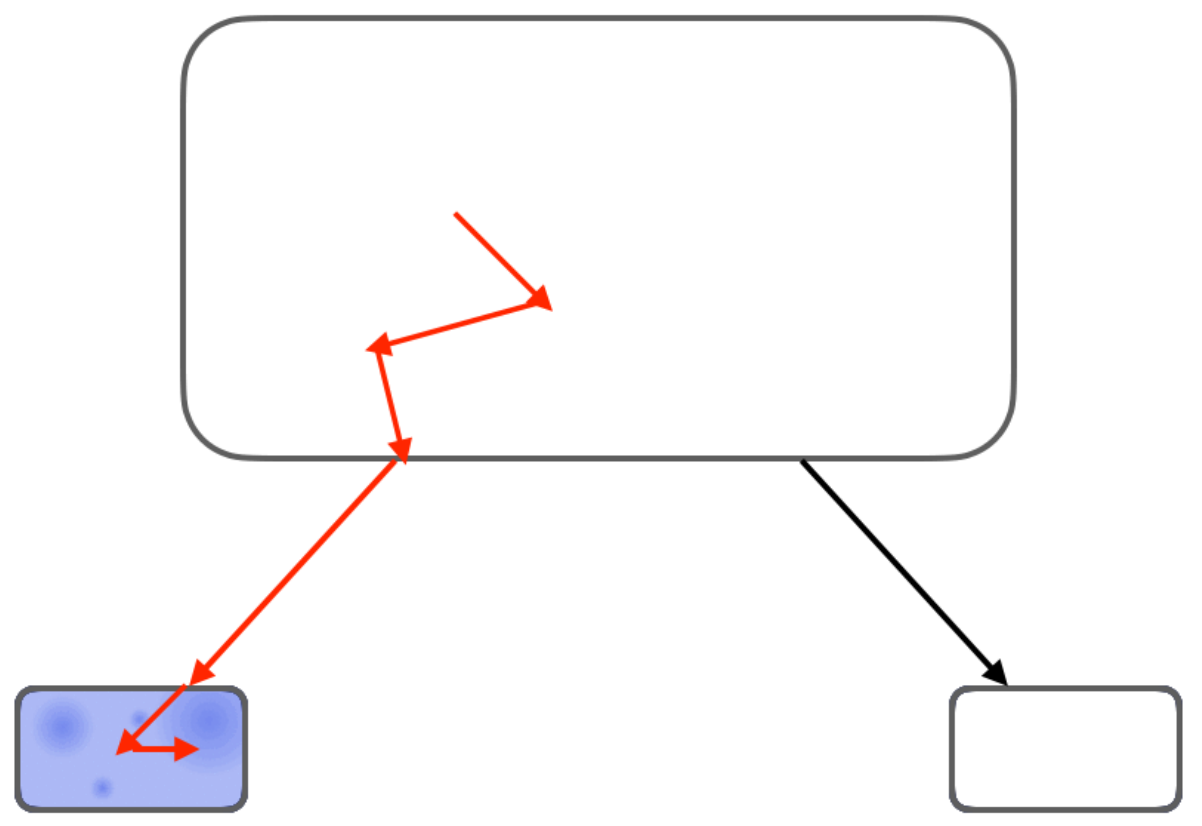
Stationary Distribution: Pros and Cons

Pros:

- Reward having in-edges from high-quality pages
- Easy to compute

Cons:

- Not defined on all graphs (only on ergodic graphs)
- Not defined for the web graph!



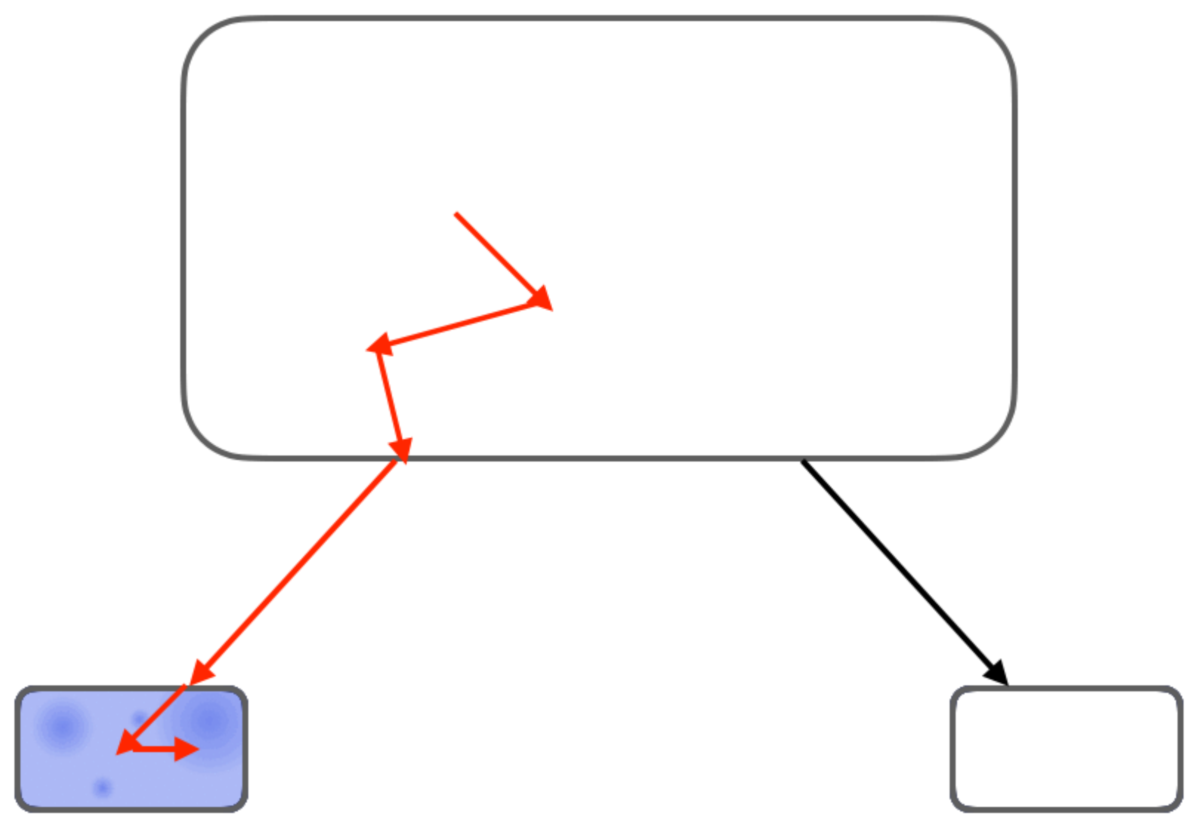
Stationary Distribution: Pros and Cons

Pros:

- Reward having in-edges from high-quality pages
- Easy to compute

Cons:

- Not defined on all graphs (only on ergodic graphs)
- Not defined for the web graph!



What does a grad student* do when their idea doesn't work?

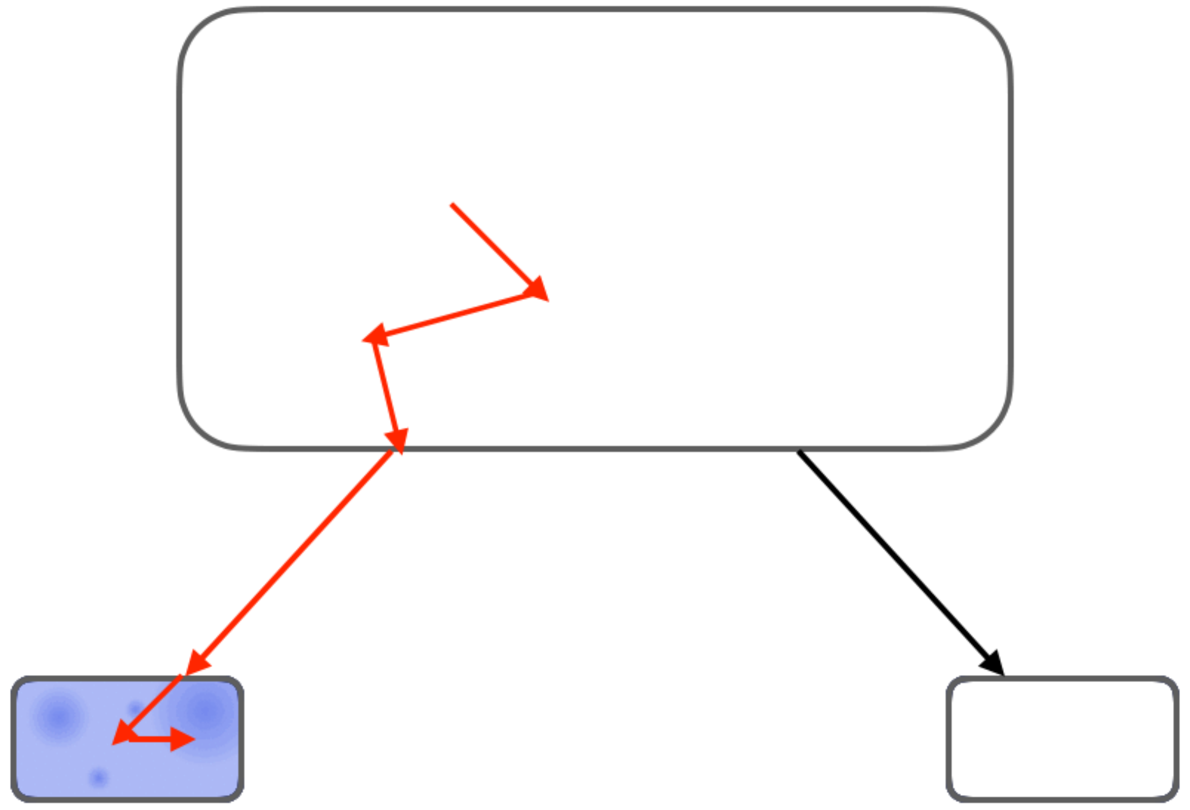
Stationary Distribution: Pros and Cons

Pros:

- Reward having in-edges from high-quality pages
- Easy to compute

Cons:

- Not defined on all graphs (only on ergodic graphs)
- Not defined for the web graph!



What does a grad student* do when their idea doesn't work?

Come up with a workaround!

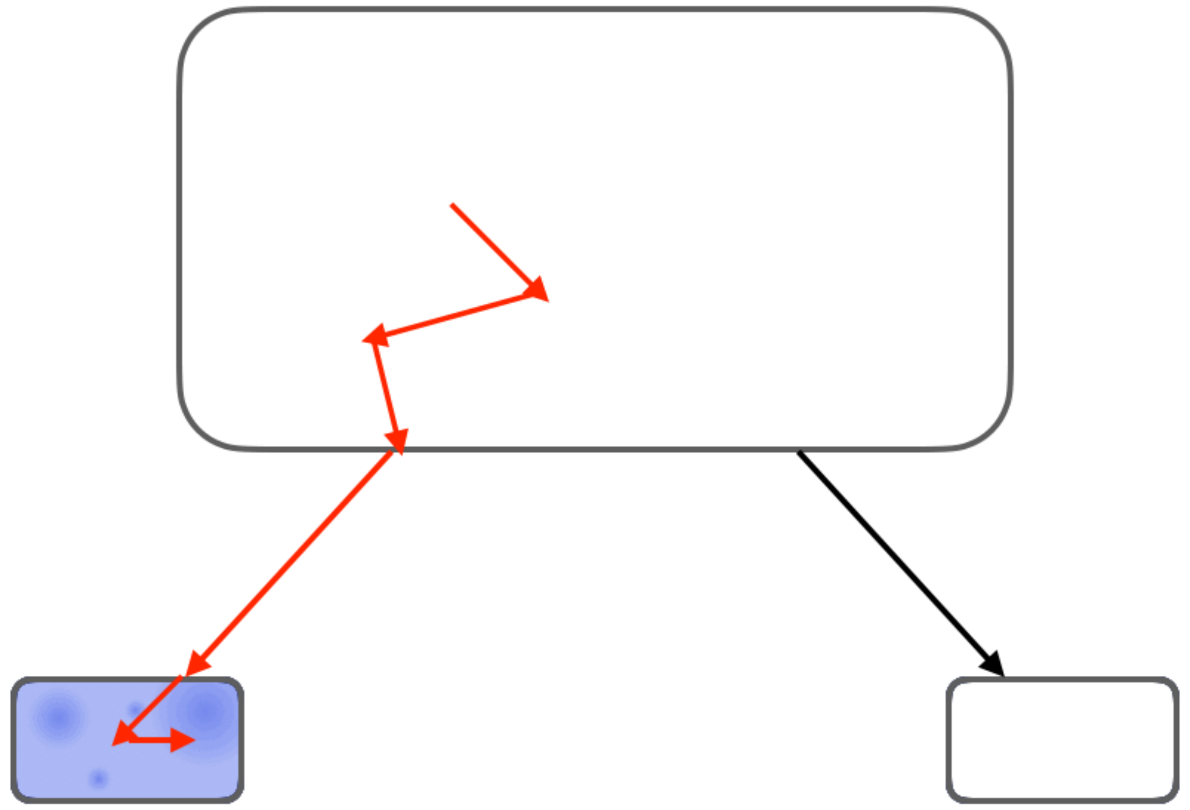
Stationary Distribution: Pros and Cons

Pros:

- Reward having in-edges from high-quality pages
- Easy to compute

Cons:

- Not defined on all graphs (only on ergodic graphs)
- Not defined for the web graph!



What does a grad student* do when their idea doesn’t work?

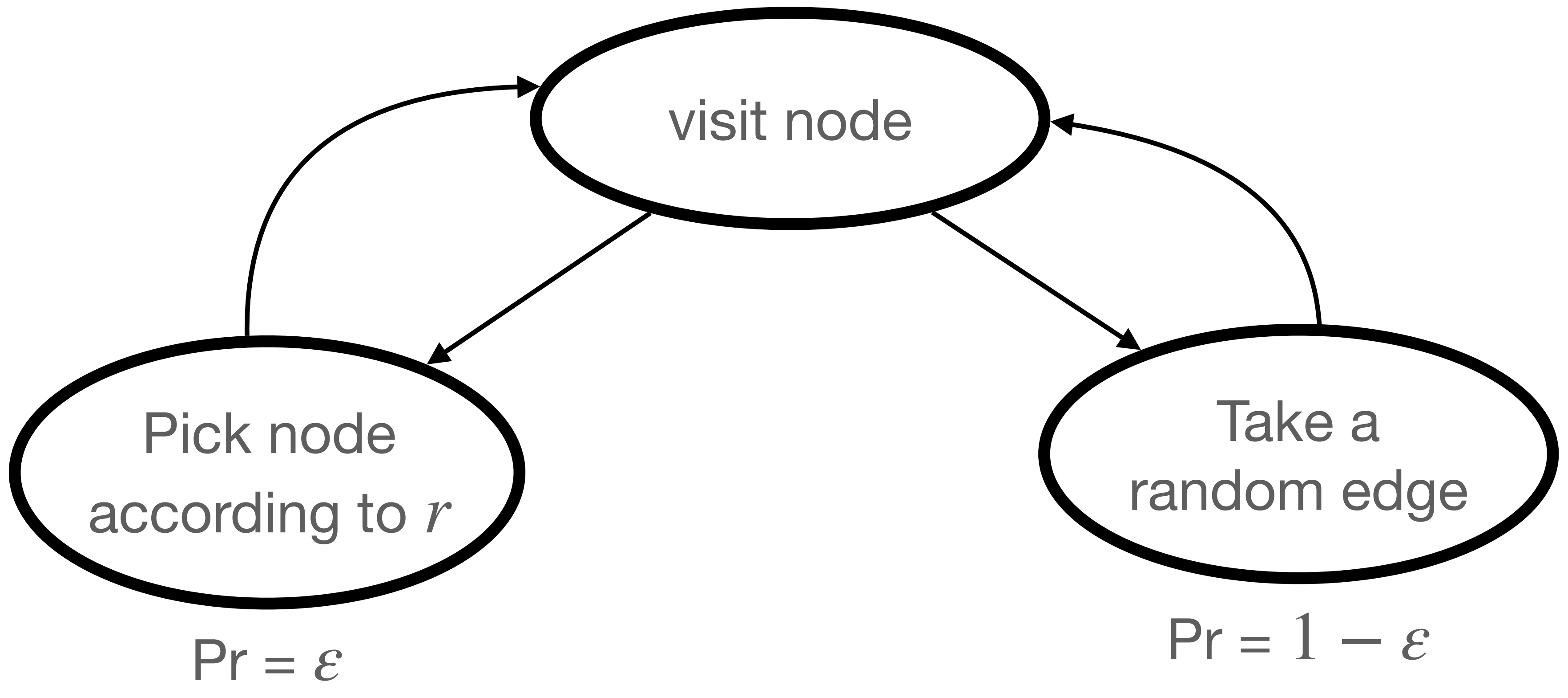
Come up with a workaround!

*NB: Grad students = Brin and Page

PageRank is the stationary distribution of a slightly different random walk

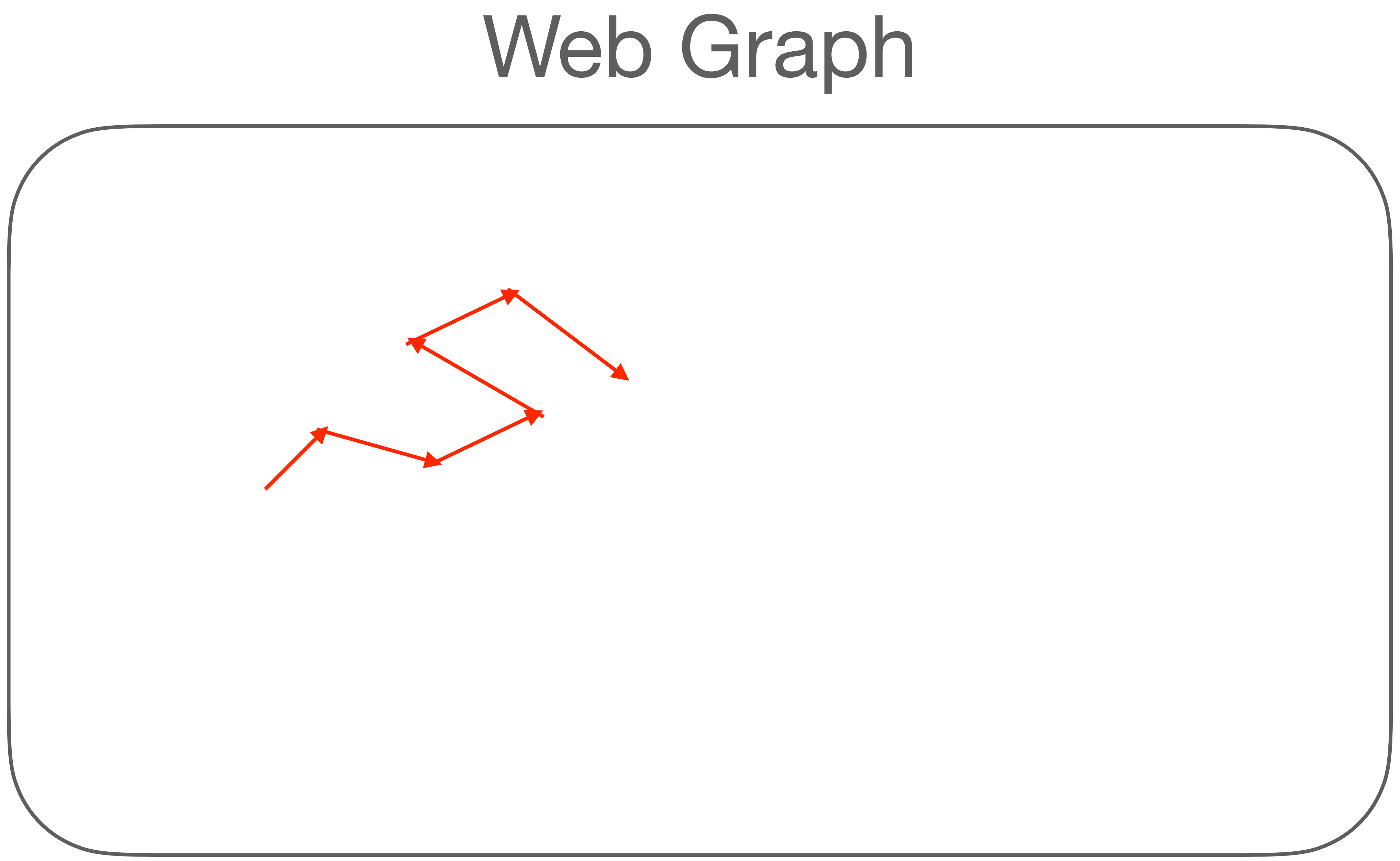
- Let *reset vector* r be a distribution over the nodes
- In the PageRank paper, $r = u$, the uniform distribution over the nodes

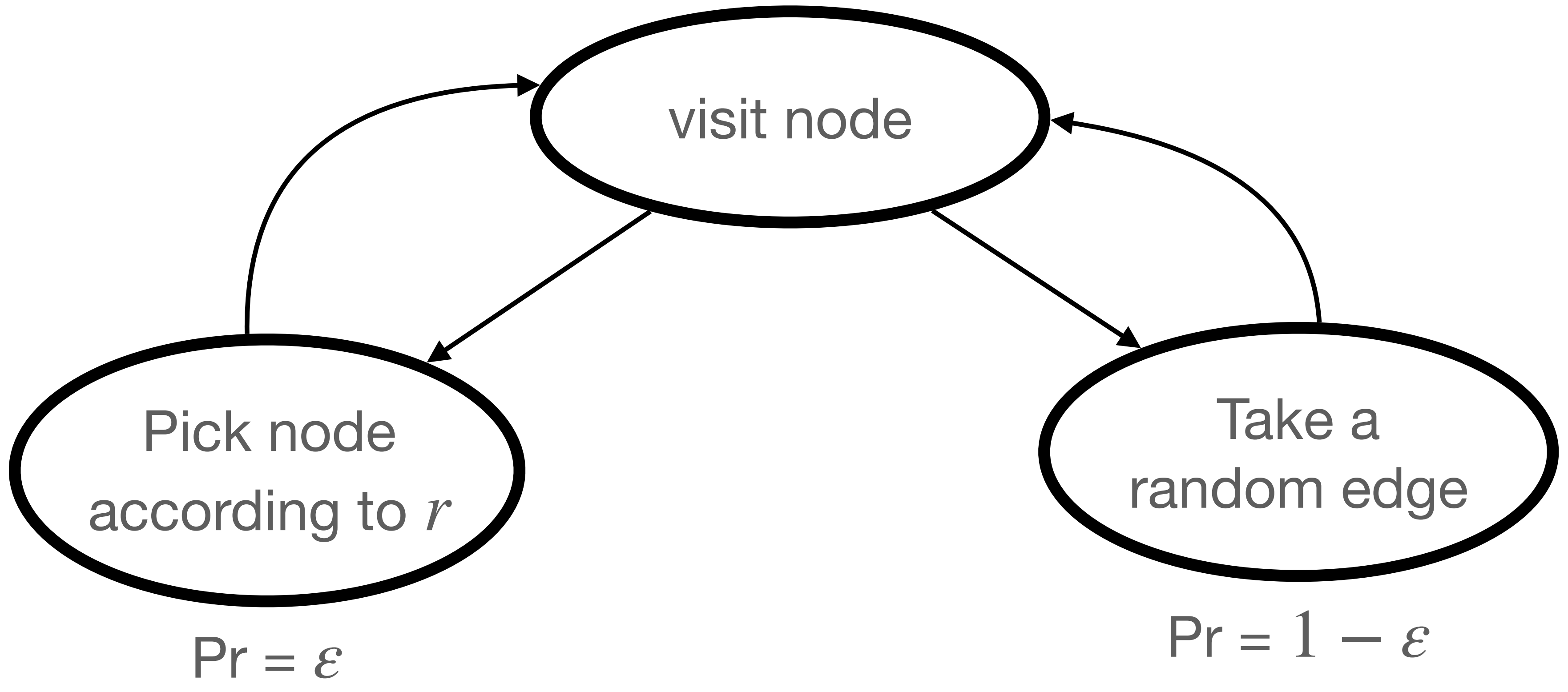
Web Graph



PageRank is the stationary distribution of a slightly different random walk

- Let *reset vector* r be a distribution over the nodes
- In the PageRank paper, $r = u$, the uniform distribution over the nodes

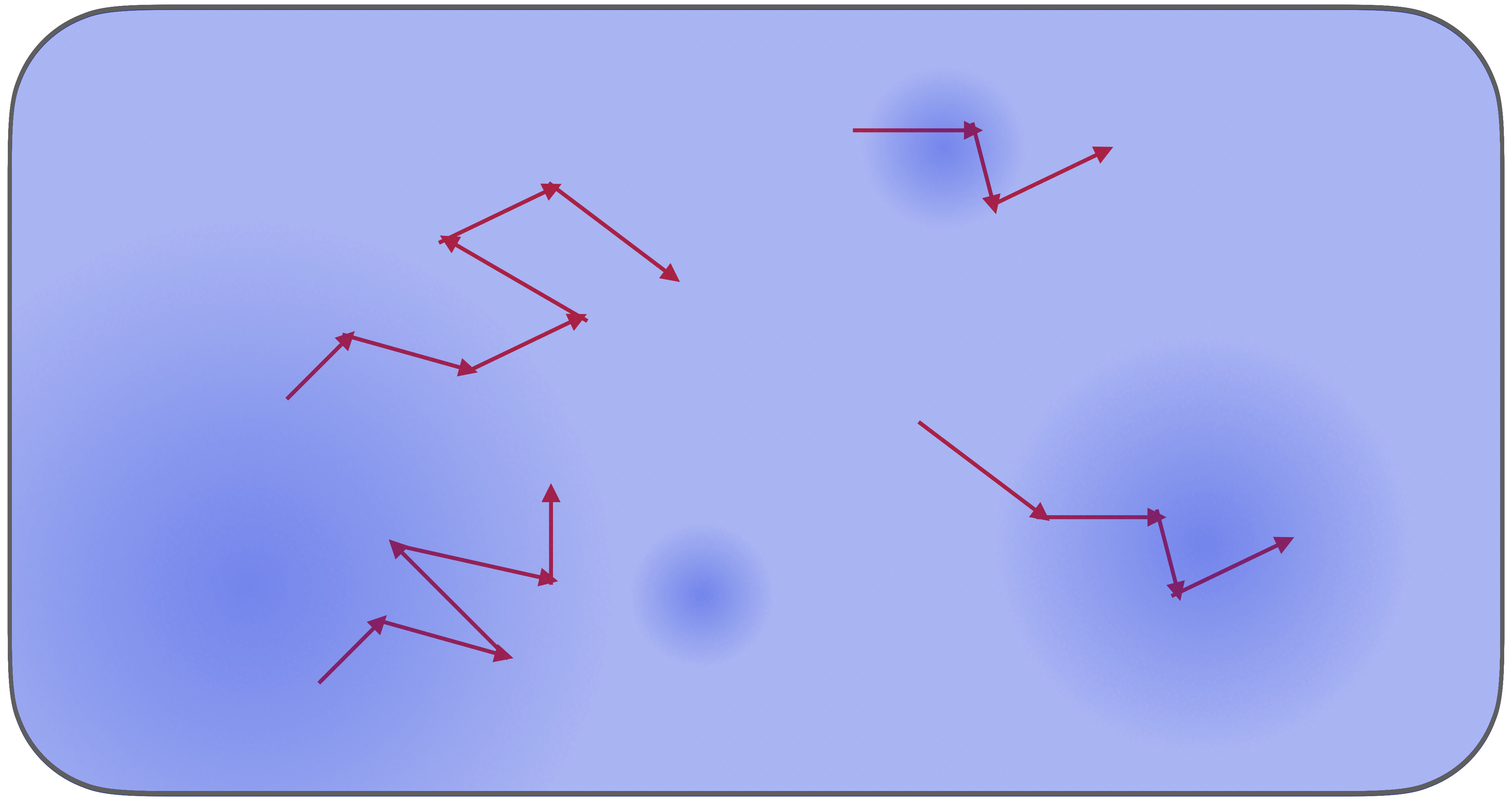


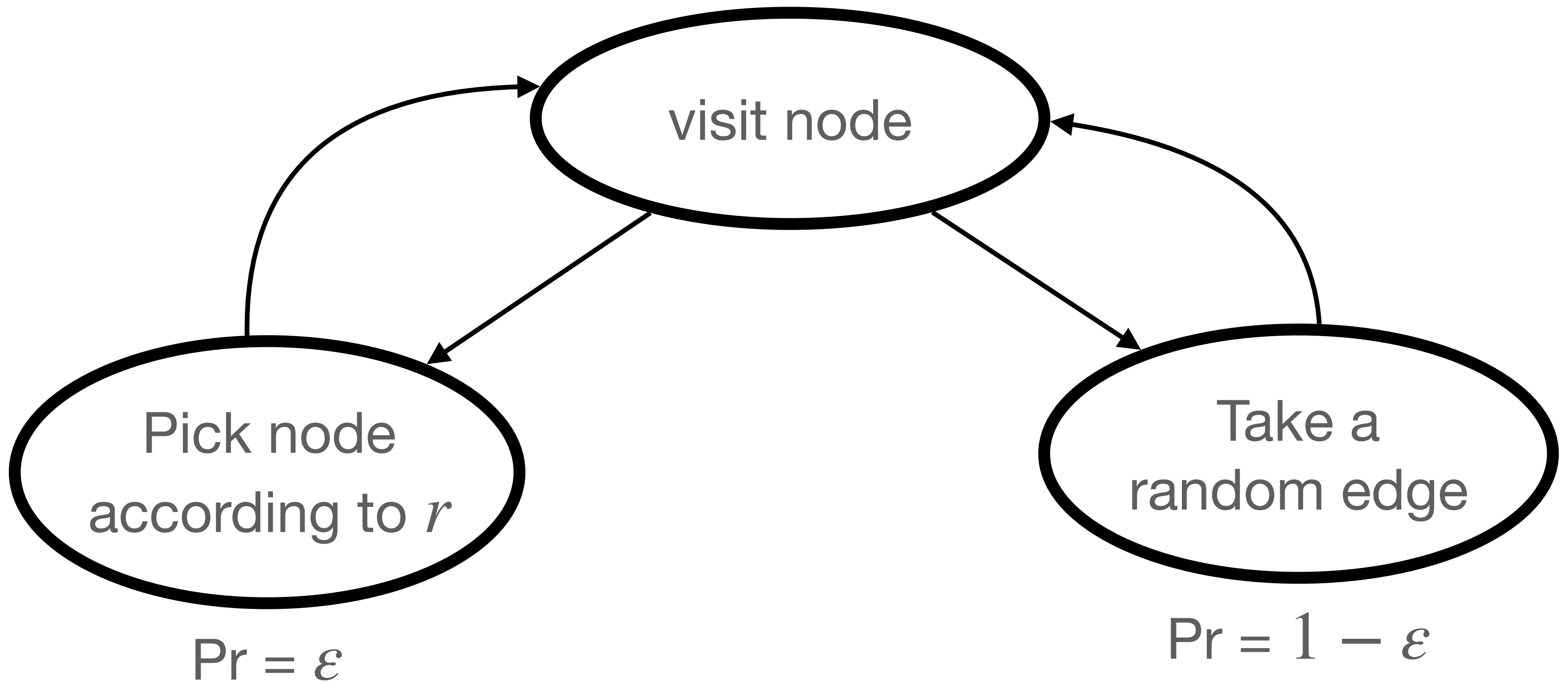


PageRank is the stationary distribution of a slightly different random walk

- Let *reset vector* r be a distribution over the nodes
- In the PageRank paper, $r = u$, the uniform distribution over the nodes

Web Graph

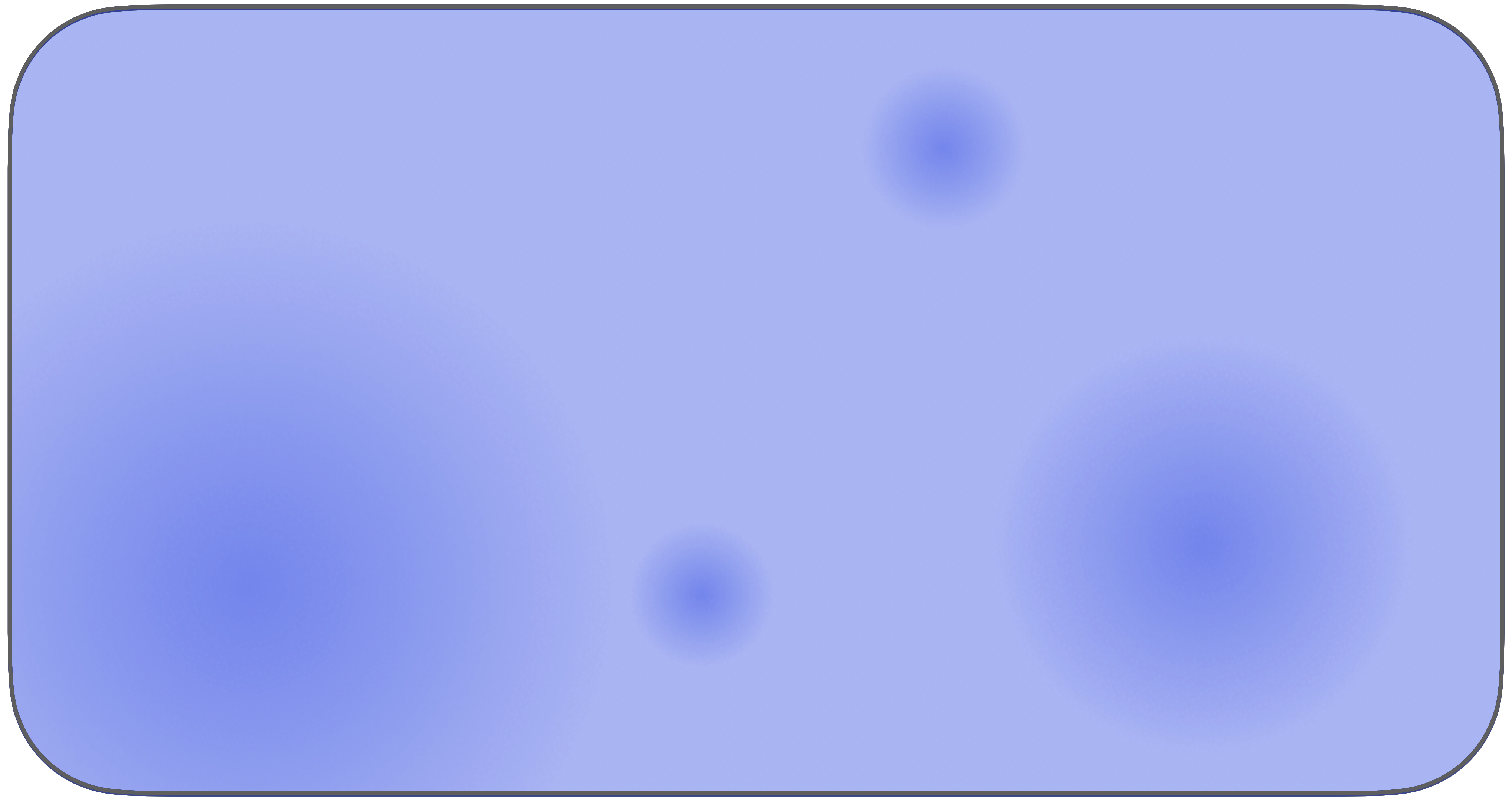




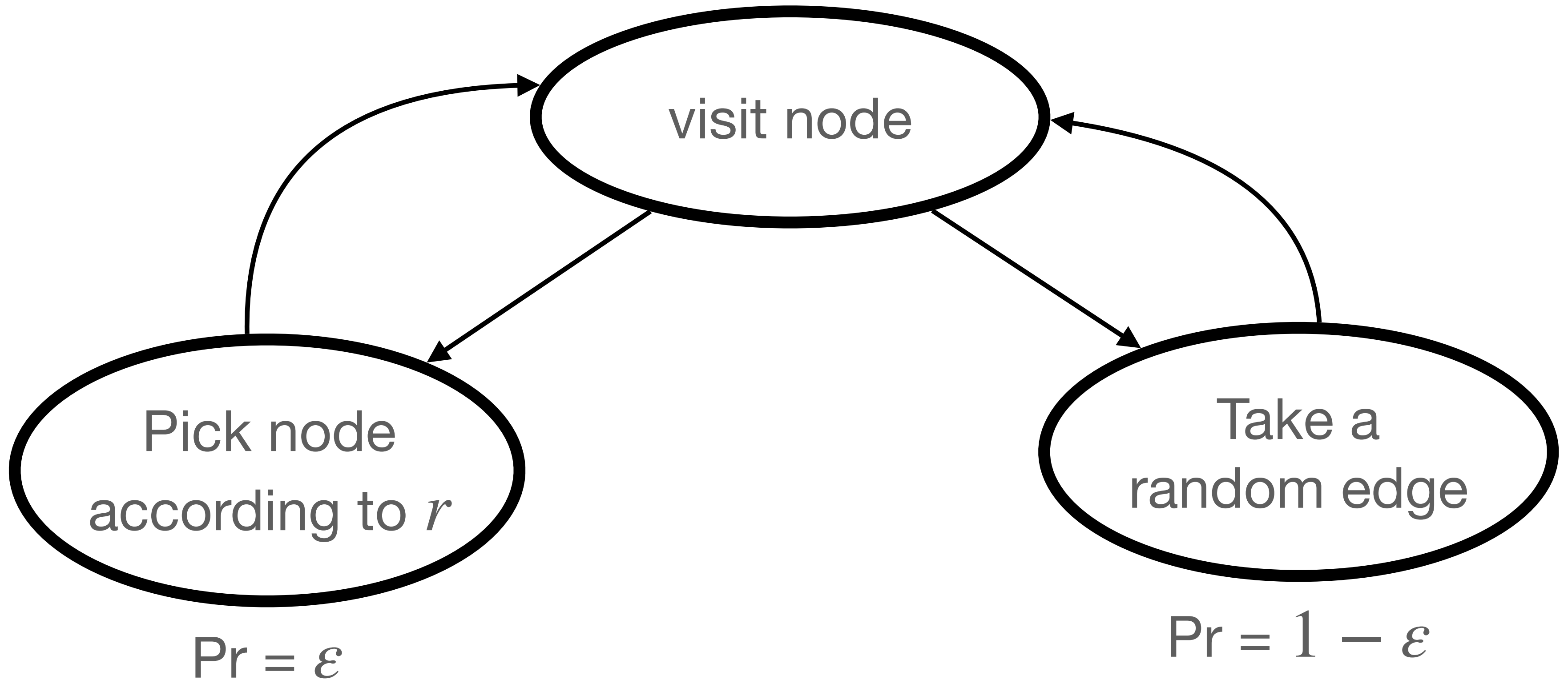
PageRank is the stationary distribution of a slightly different random walk

- Let *reset vector* r be a distribution over the nodes
- In the PageRank paper, $r = u$, the uniform distribution over the nodes

Web Graph



PageRank = stationary distribution



PageRank: the workaround that saved the stationary distribution idea

PageRank is the stationary distribution of a slightly different random walk

- Let *reset vector* r be a distribution over the nodes
- In the PageRank paper, $r = u$, the uniform distribution over the nodes

PageRank is defined for every graph! (unless $\varepsilon = 0$)

PageRank: the workaround that saved the stationary distribution idea

PageRank is the stationary distribution of a slightly different random walk

- Let *reset vector* r be a distribution over the nodes
- In the PageRank paper, $r = u$, the uniform distribution over the nodes

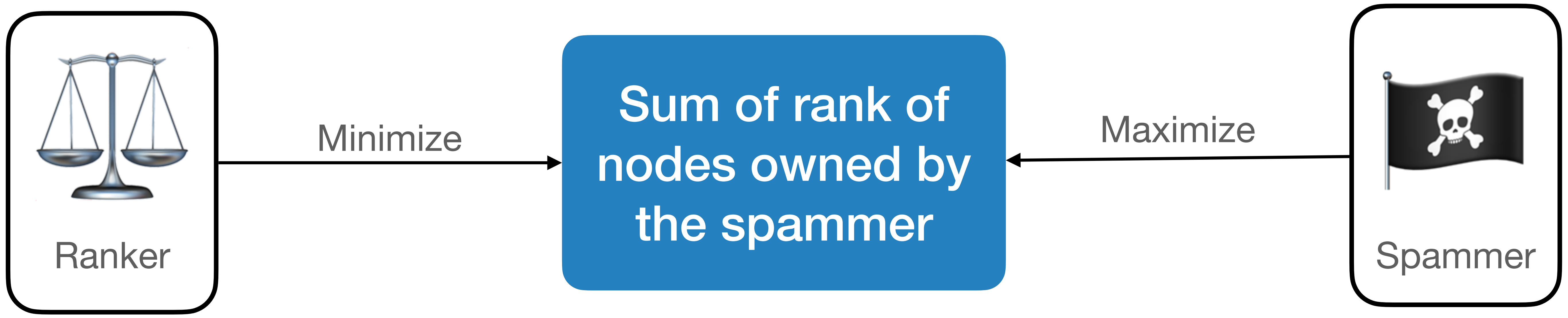
PageRank is defined for every graph! (unless $\varepsilon = 0$)

Types of PageRank:

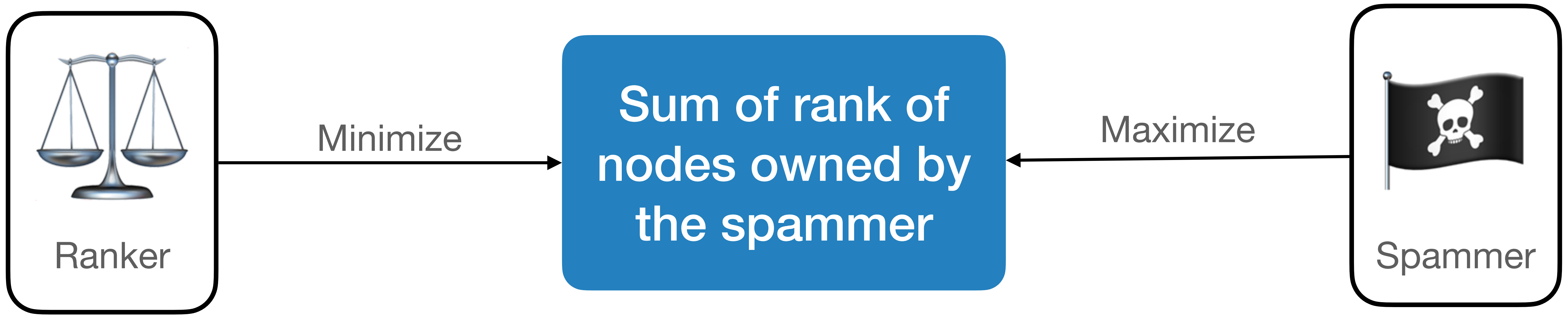
- When $r = u$, it's called the *Uniform PageRank* (UPR)
- When $r[c] = 1$, for some *center* c, it's called *Personalized PageRank* (PPR_c)

Contribution: Spamming Game

What can the Ranker and Spammer do?

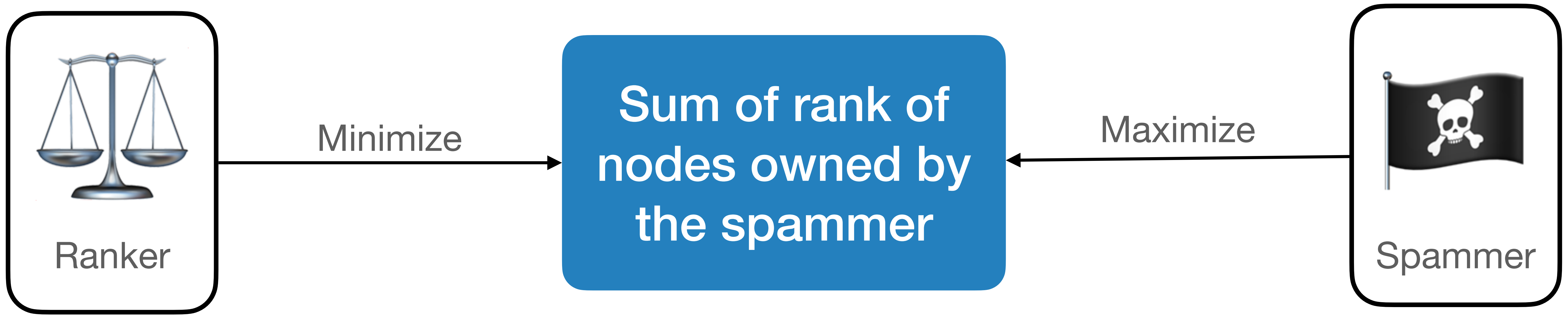


What can the Ranker and Spammer do?



Select a Ranking Function
(within reason)

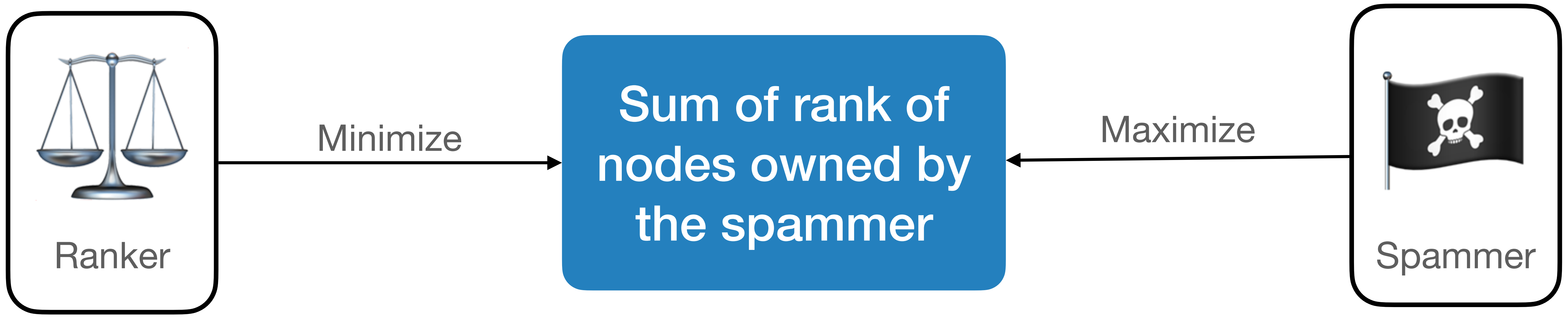
What can the Ranker and Spammer do?



Select a Ranking Function
(within reason)

Create new nodes (for free)

What can the Ranker and Spammer do?

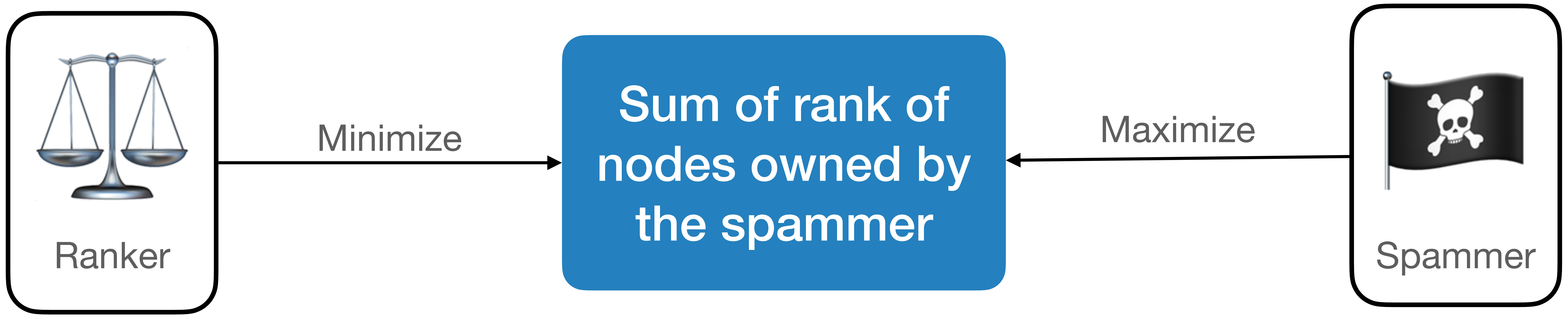


Select a Ranking Function
(within reason)

Create new nodes (for free)

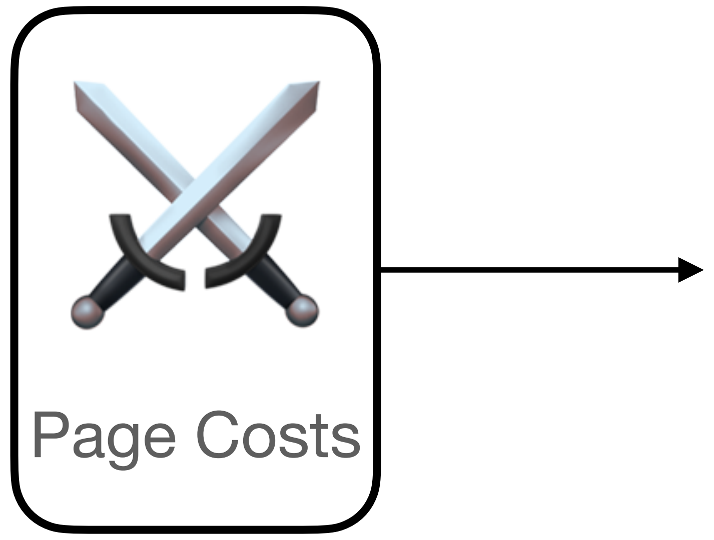
Acquire previously existing nodes at a cost

What can the Ranker and Spammer do?



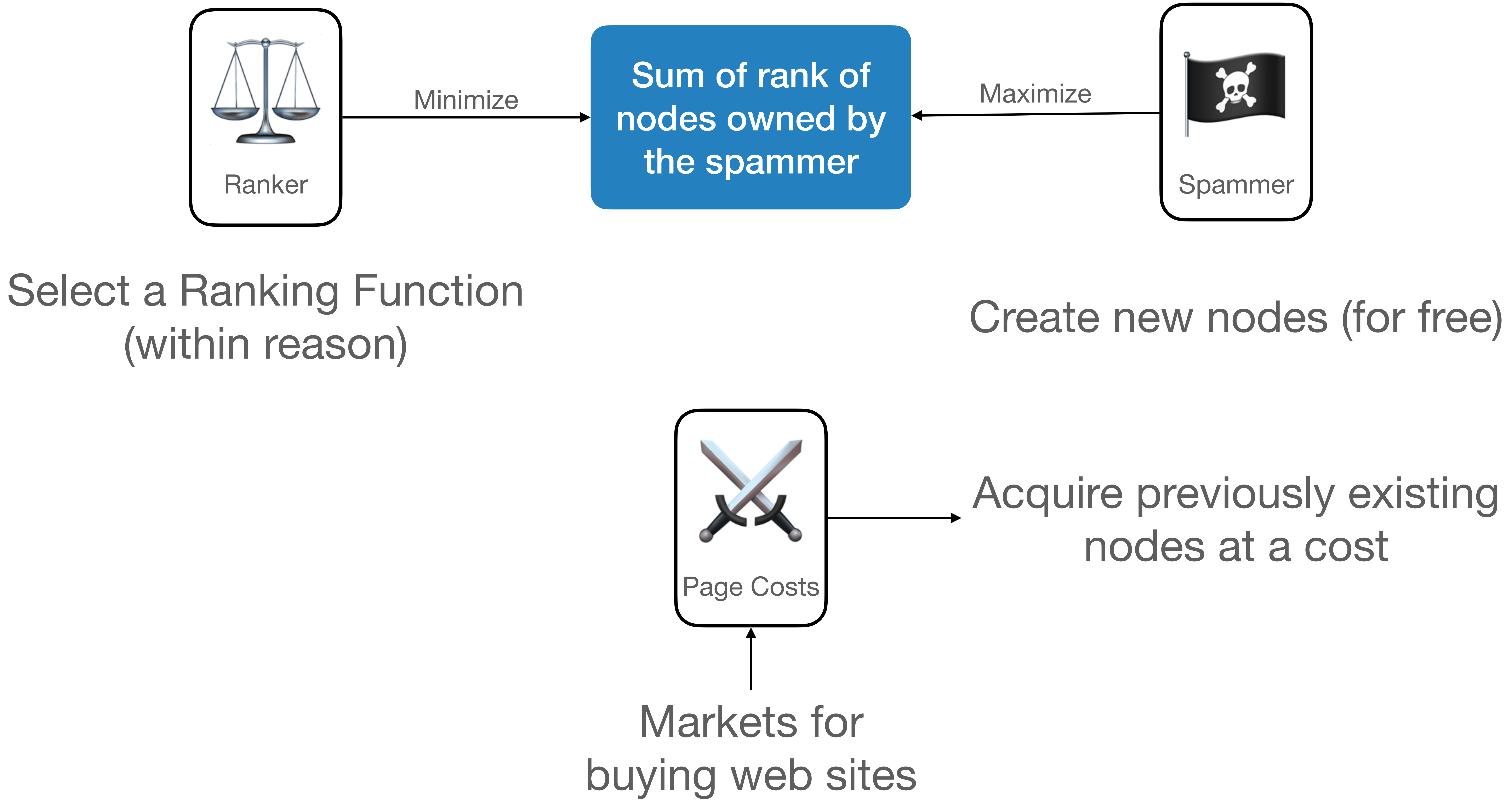
Select a Ranking Function
(within reason)

Create new nodes (for free)

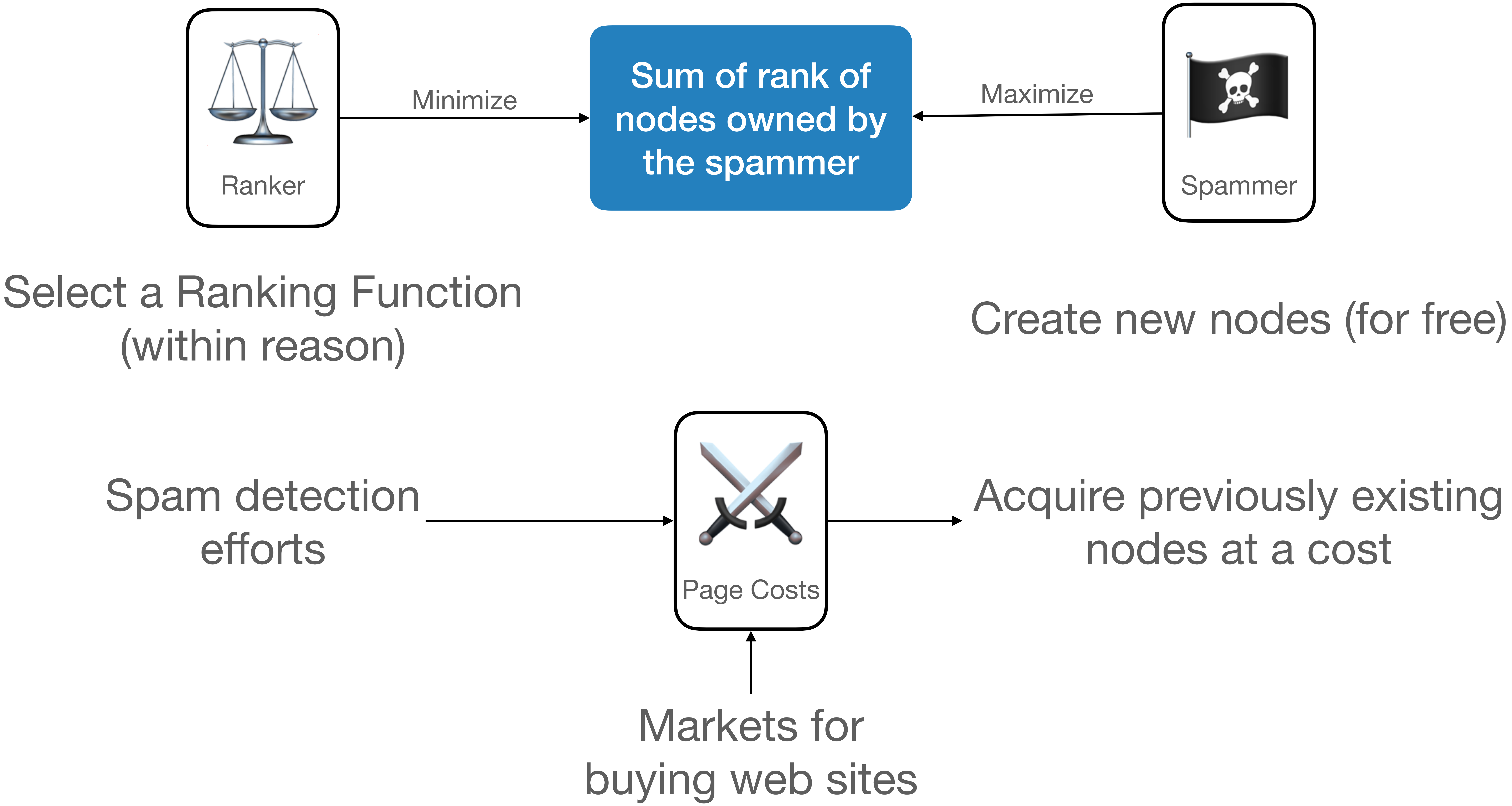


Acquire previously existing nodes at a cost

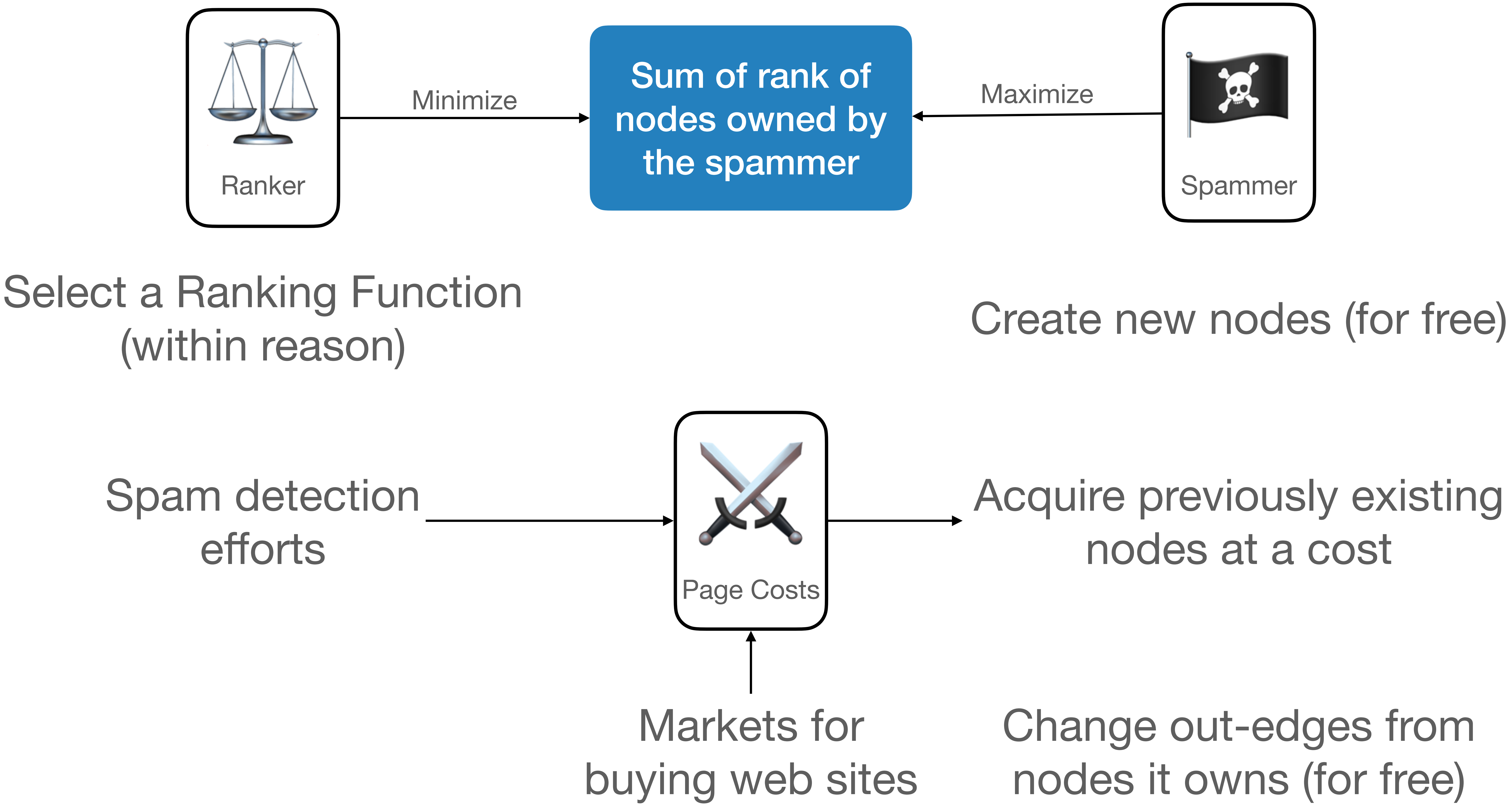
What can the Ranker and Spammer do?



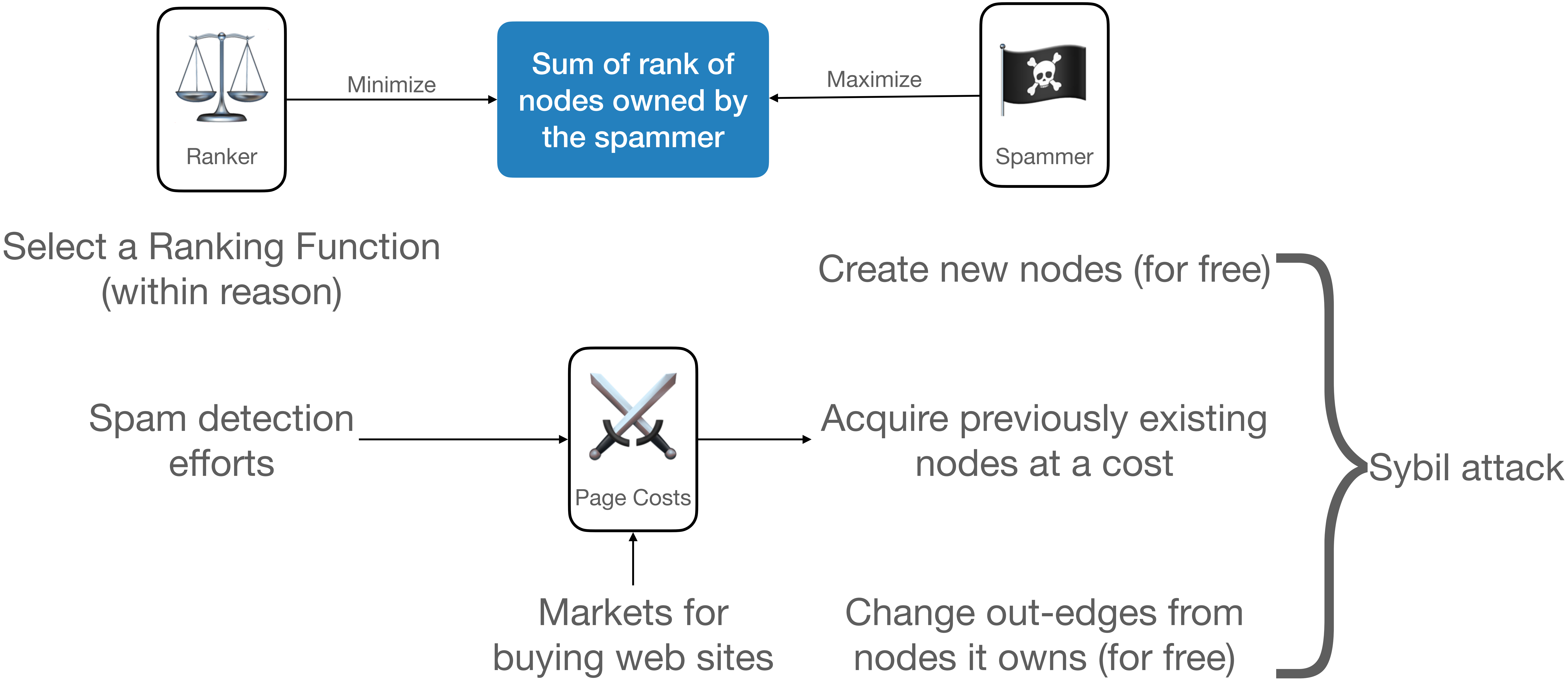
What can the Ranker and Spammer do?



What can the Ranker and Spammer do?



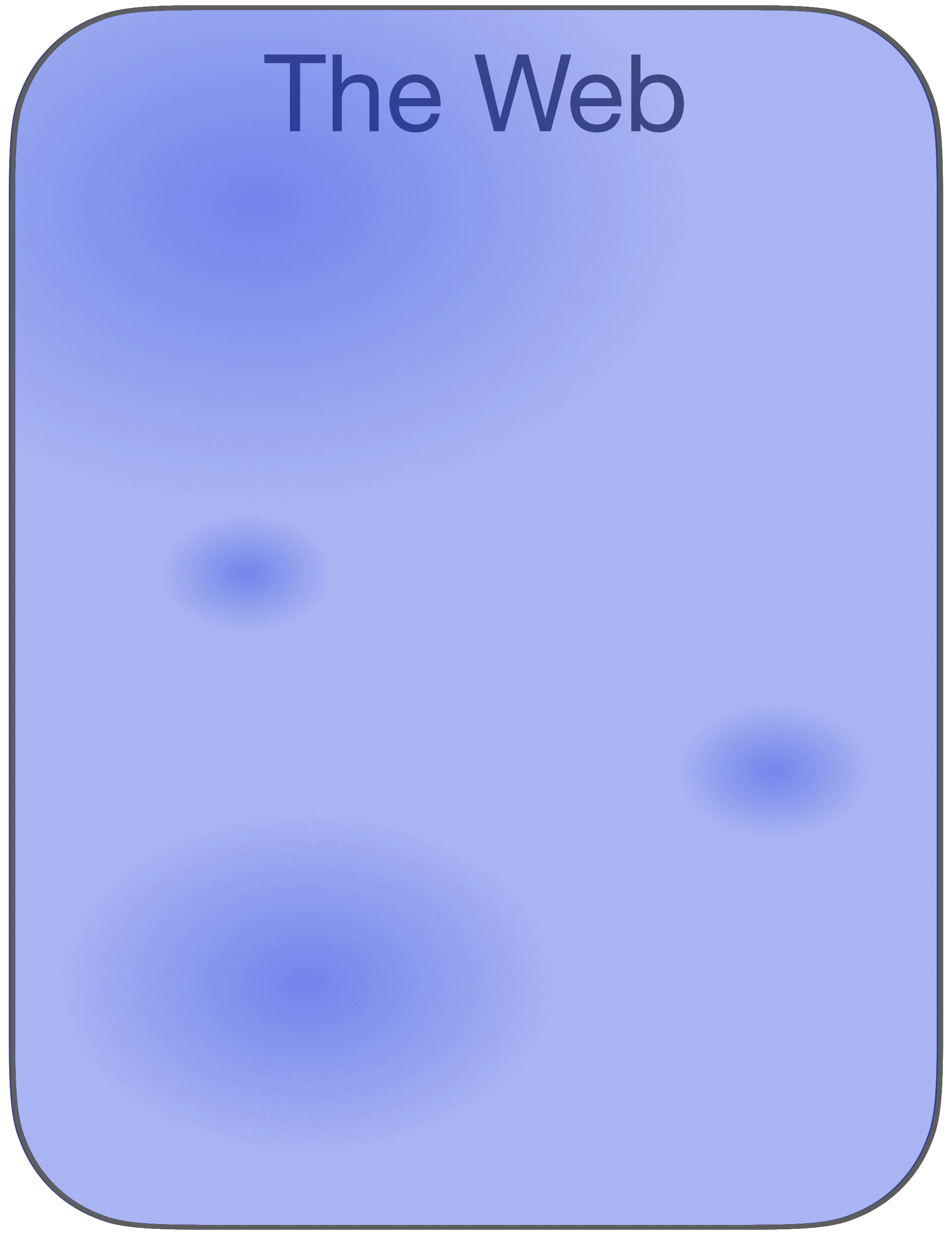
What can the Ranker and Spammer do?



UPR Spamming

Example: Uniform PageRank

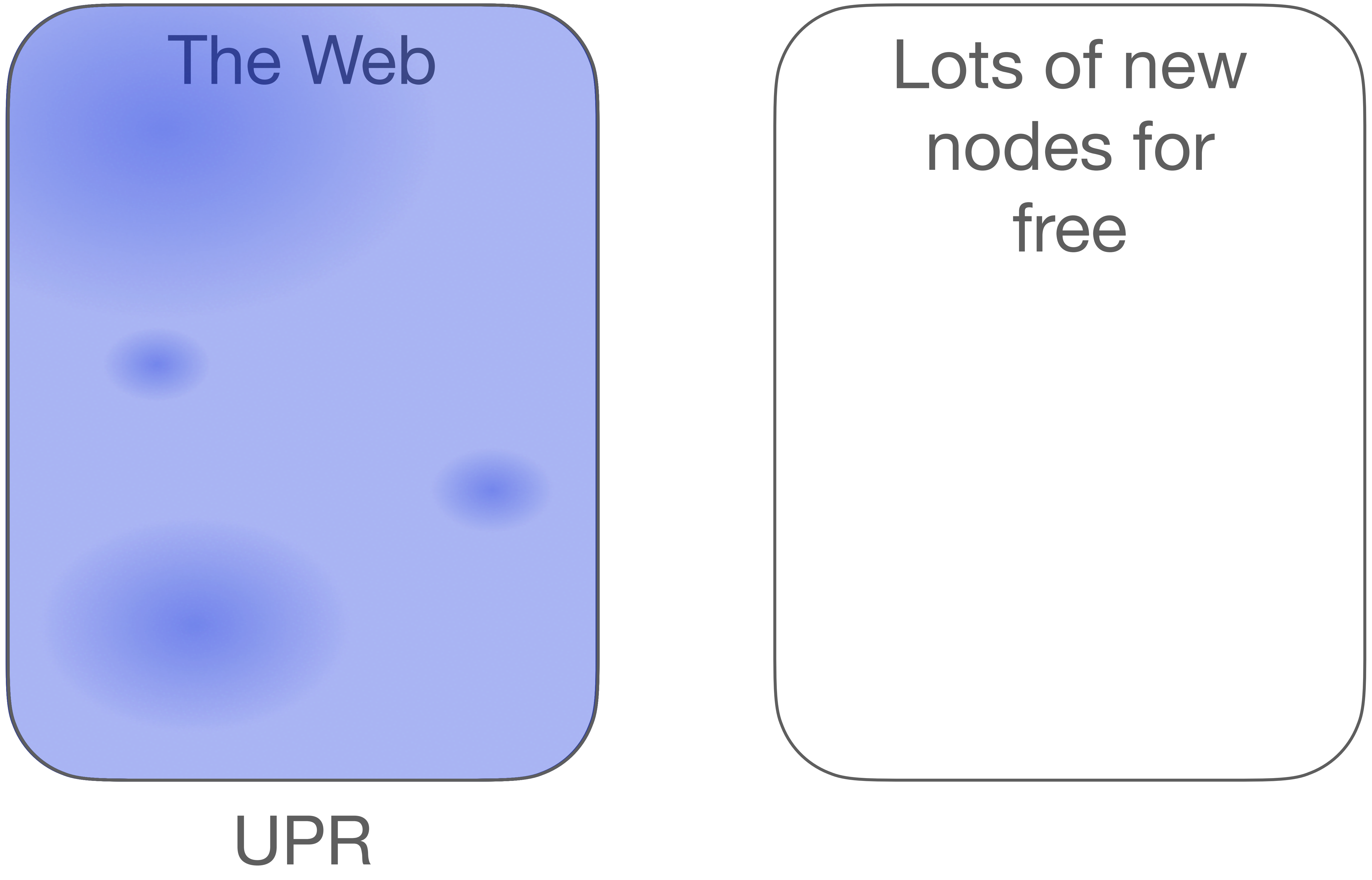
If ranker is using UPR, what should the spammer do?



UPR

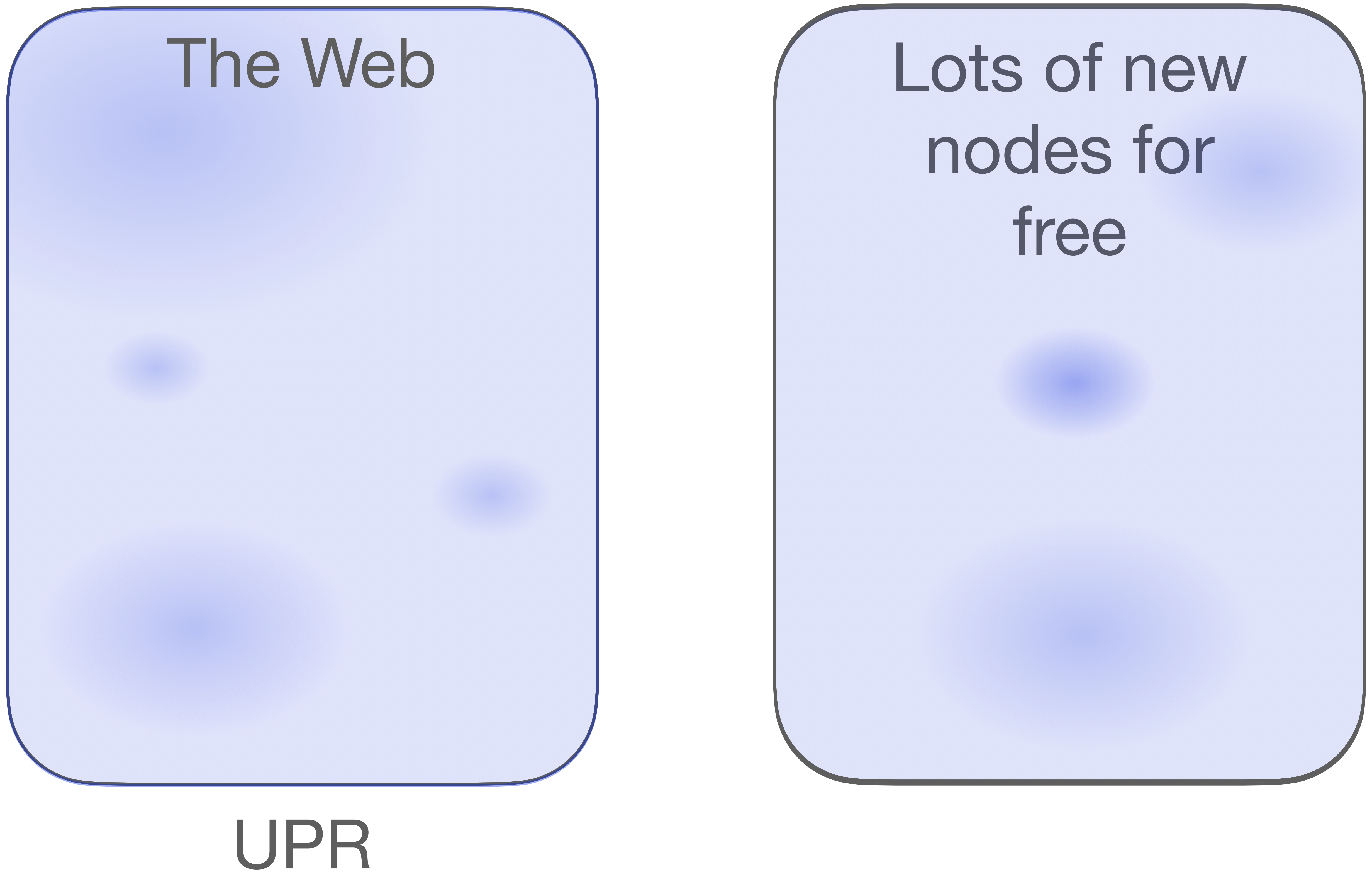
Example: Uniform PageRank

If ranker is using UPR, what should the spammer do?



Example: Uniform PageRank

If ranker is using UPR, what should the spammer do?

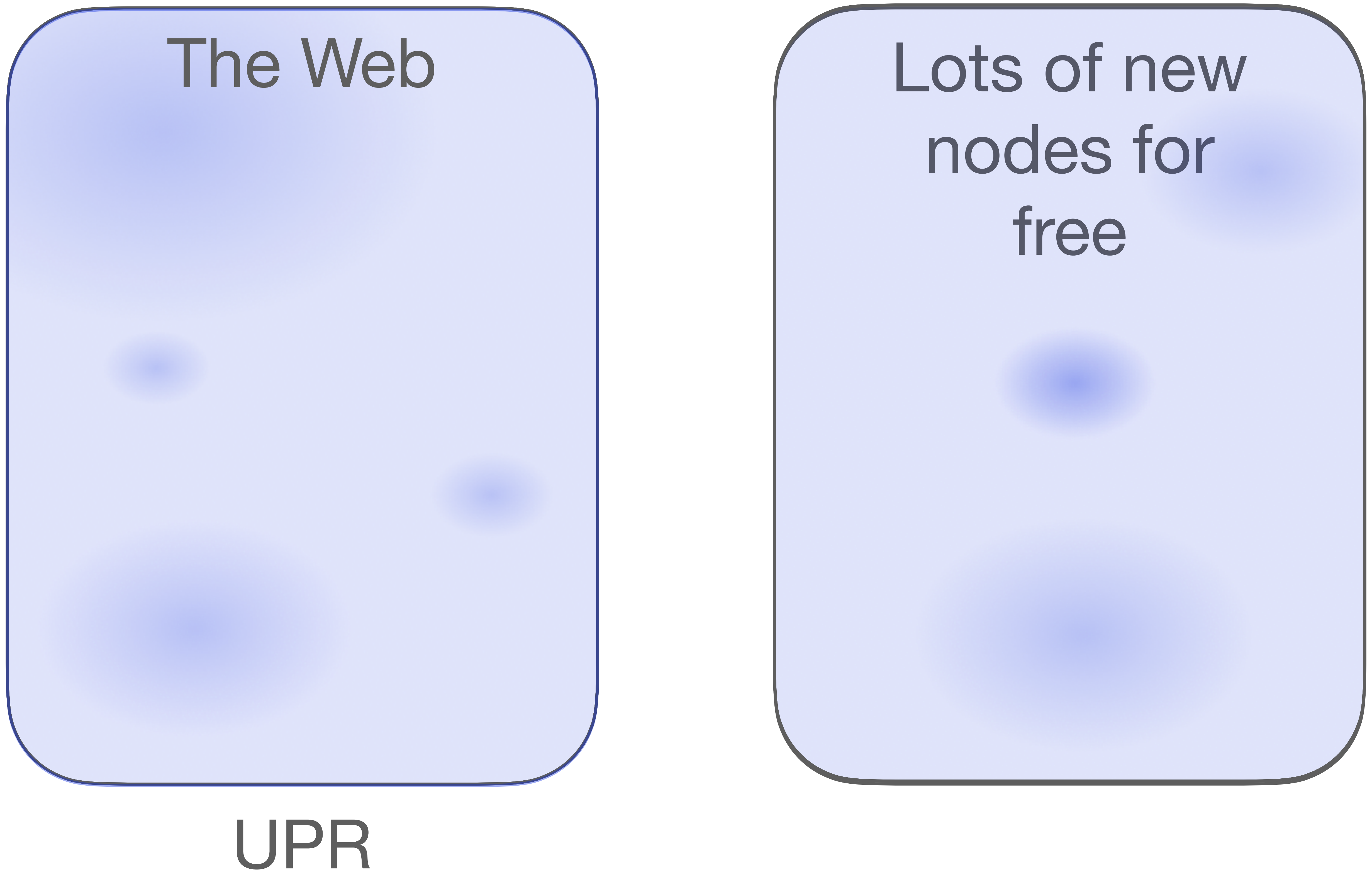


PageRank will reset to new nodes quite often

So spammer can acquire as much rank as they want

With no expenditure!

If ranker is using UPR, what should the spammer do?



PageRank will reset to new nodes quite often

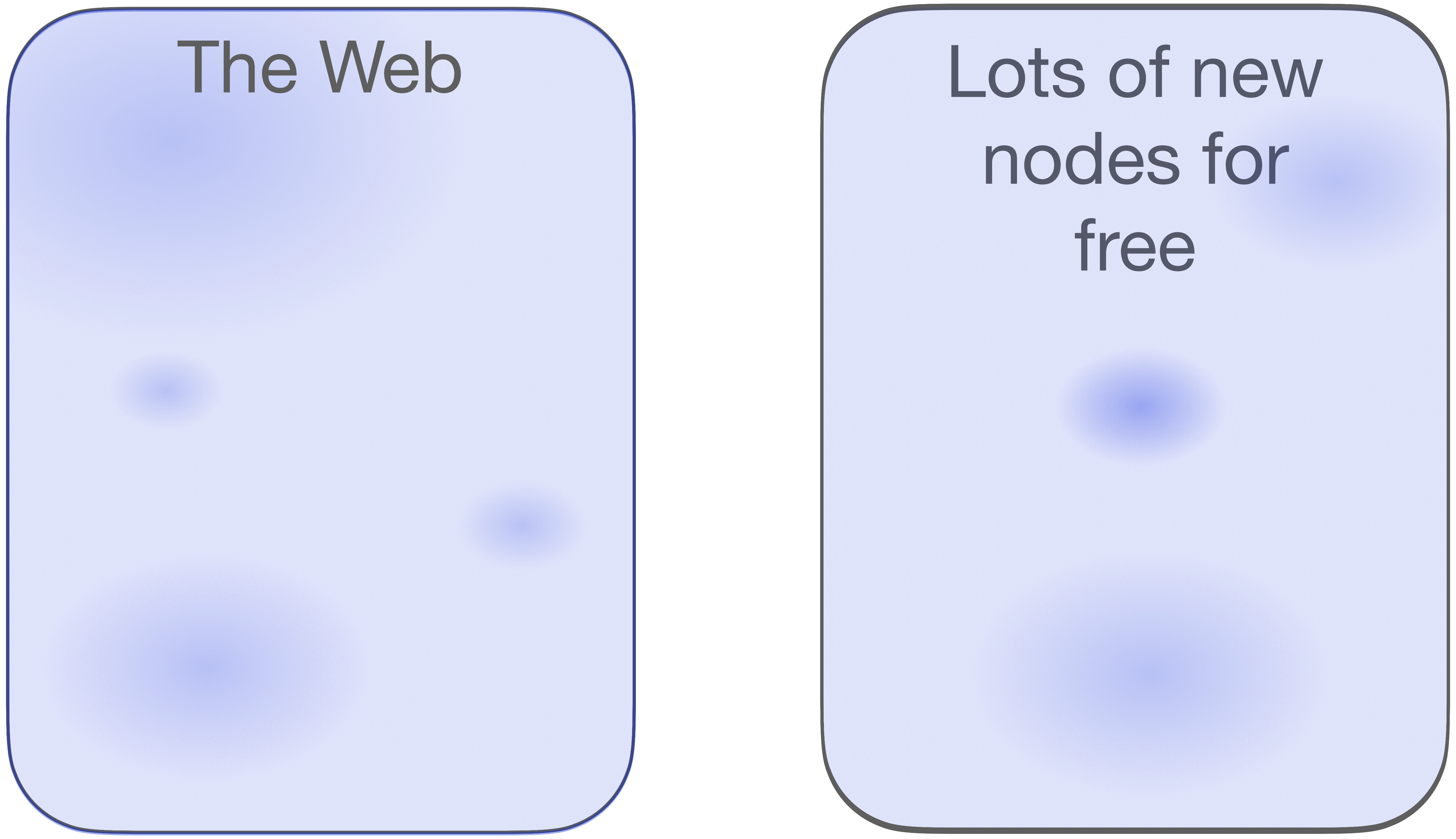
So spammer can acquire as much rank as they want

With no expenditure!

So total captured rank approaches 1.

Example: Uniform PageRank

If ranker is using UPR, what should the spammer do?



UPR

PageRank will reset to new nodes quite often

So spammer can acquire as much rank as they want

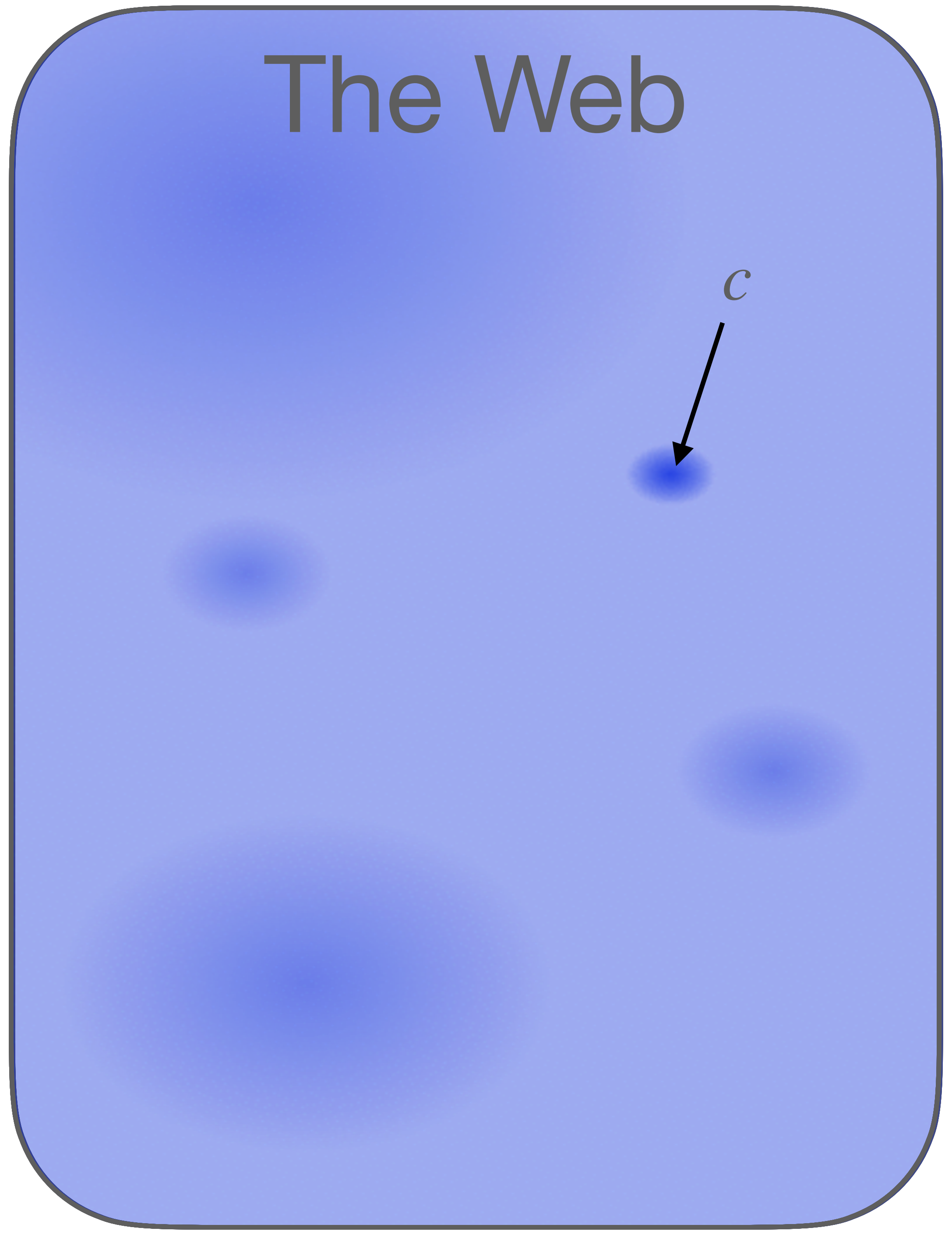
With no expenditure!

So total captured rank approaches 1.

This trivial vulnerability of UPR is why it was never used for web ranking at Google!

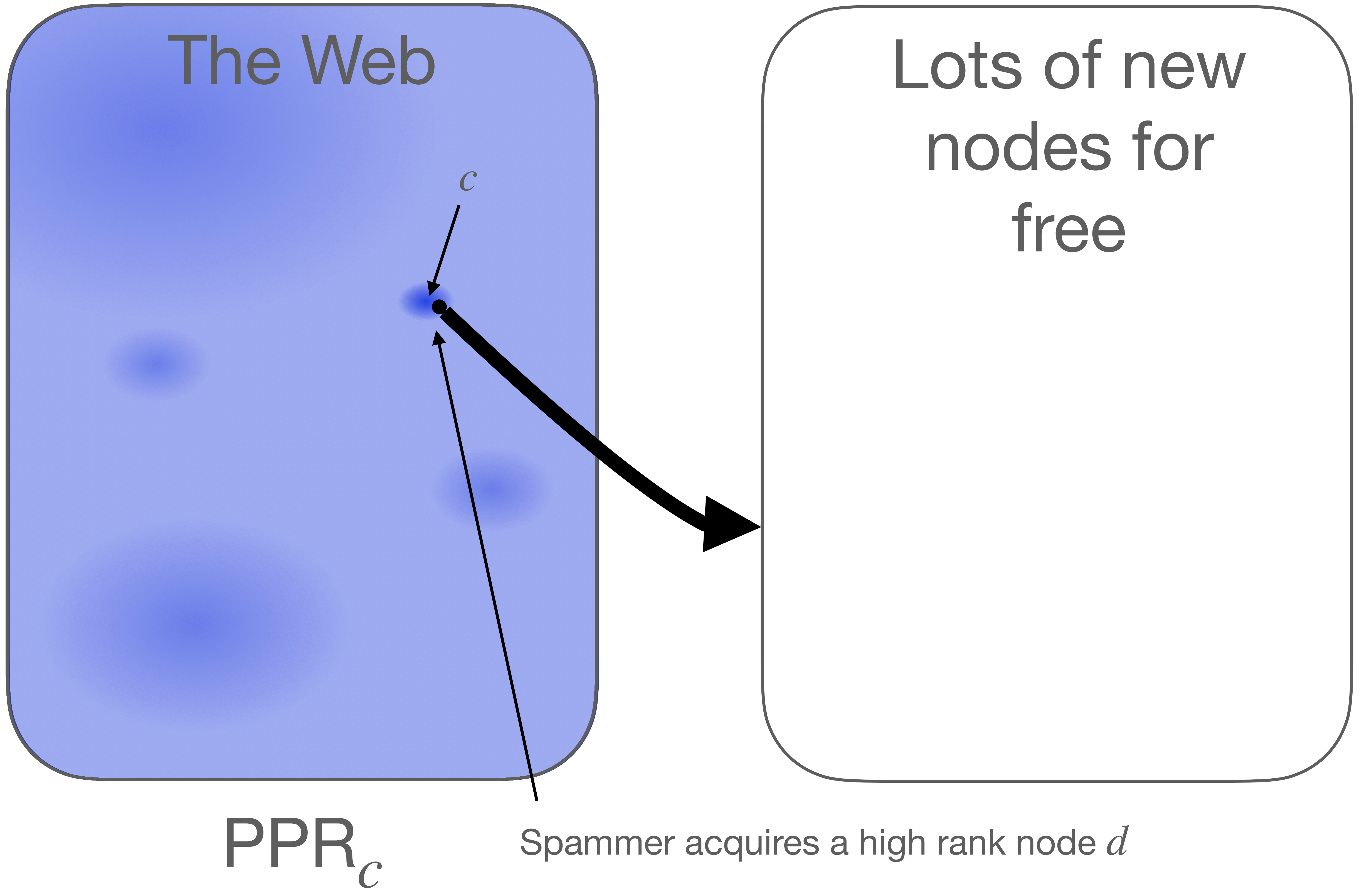
PPR Spamming

If ranker is using PPR, what should the spammer do?



PPR_c

If ranker is using PPR, what should the spammer do?

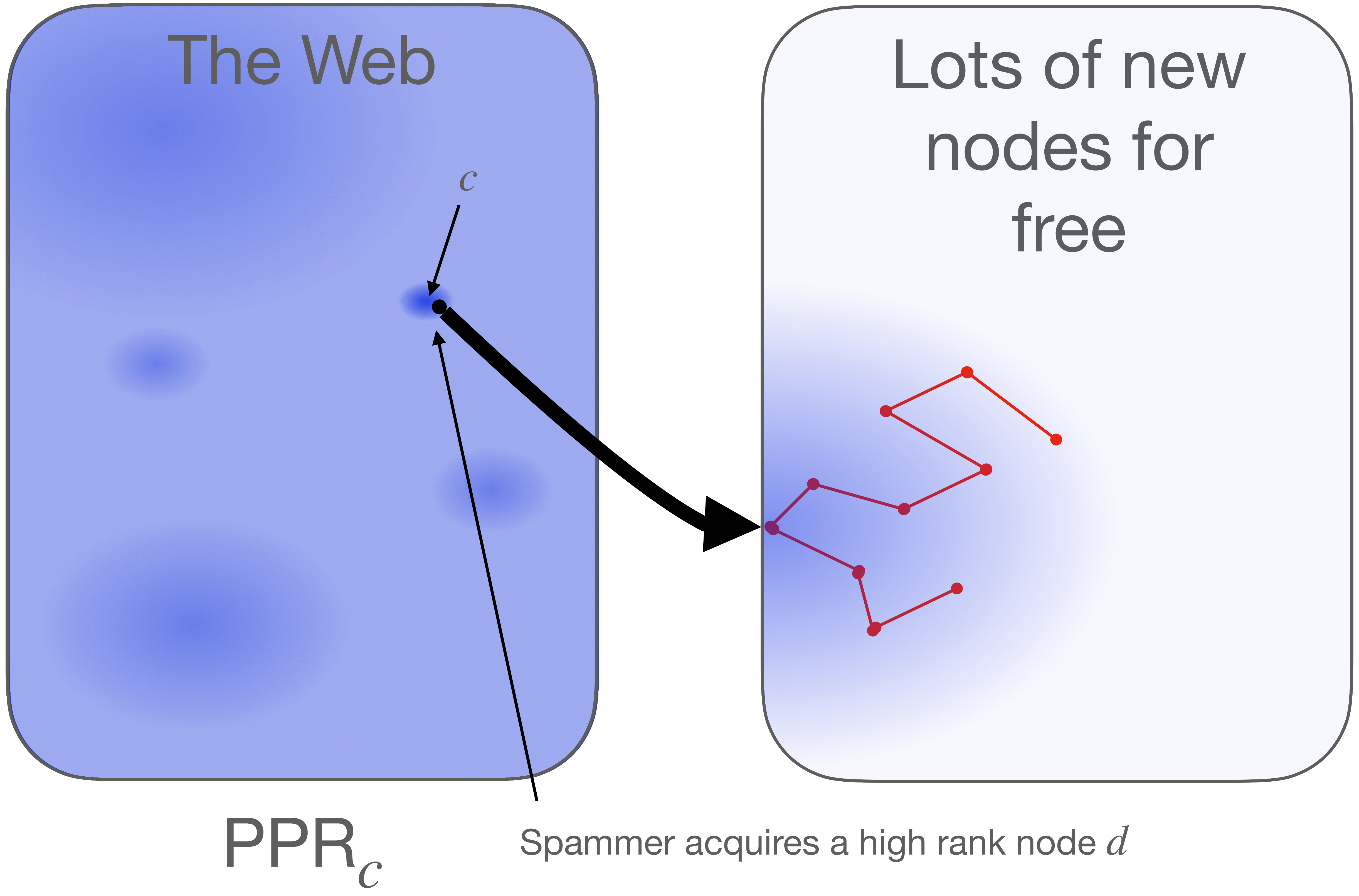


Spammer must acquire a high rank node

It can then divert random walks into it's own new nodes

Where they last for about $1/\varepsilon$ steps before reseting to c

If ranker is using PPR, what should the spammer do?

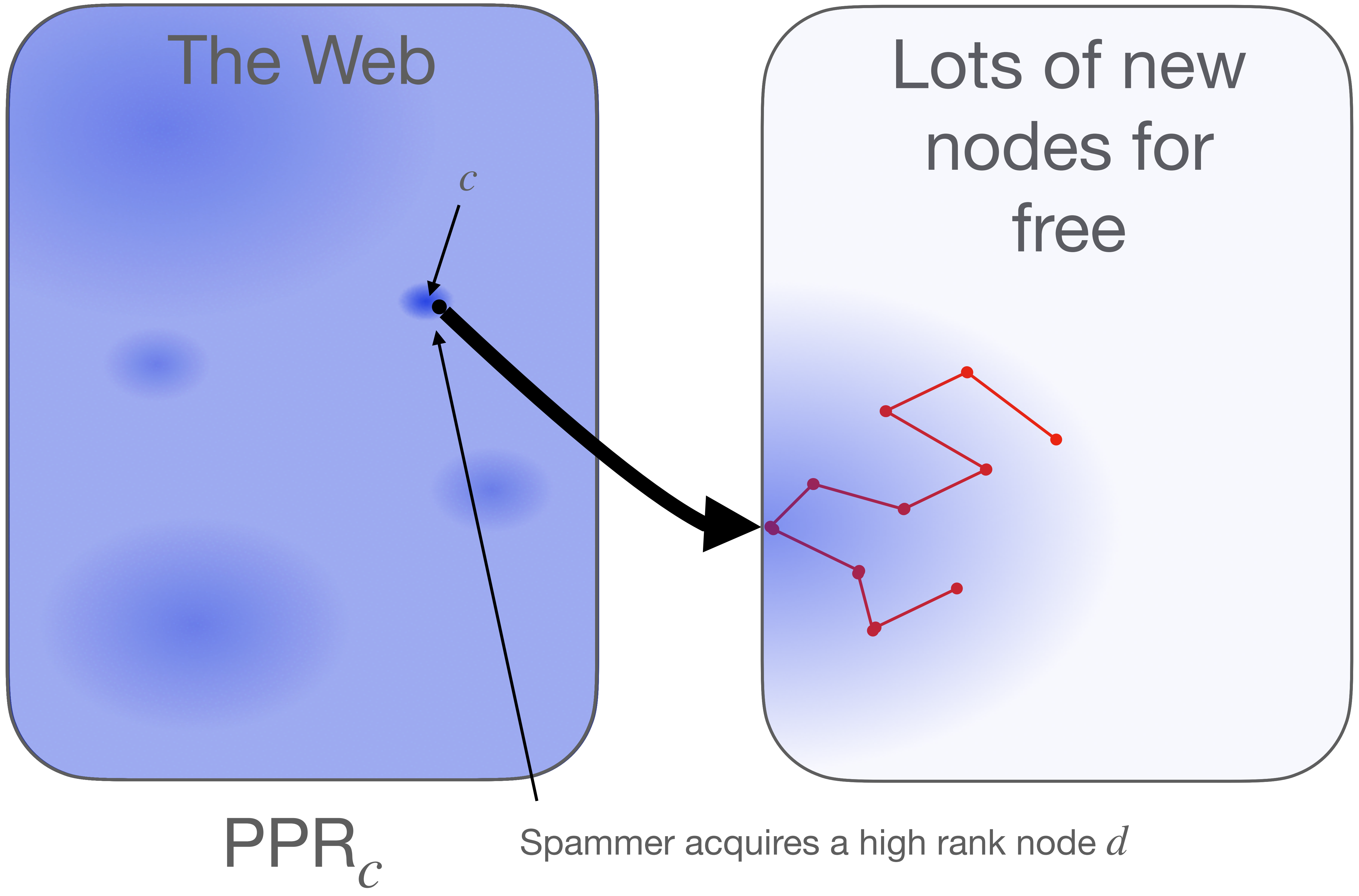


Spammer must acquire a high rank node

It can then divert random walks into it's own new nodes

Where they last for about $1/\varepsilon$ steps before reseting to c

If ranker is using PPR, what should the spammer do?



Spammer must acquire a high rank node

It can then divert random walks into it's own new nodes

Where they last for about $1/\varepsilon$ steps before reseting to c

So total captured rank is $\mathrm{PPR}_c(d)/\varepsilon$

Contribution: Definition of Spam Resistance

A ranking function is *σ-spam resistant* if, for every graph

$$\frac{\text{Spammer's Cost}}{\text{Spammer's Benefit}} \geq \sigma$$

Contribution: Definition of Spam Resistance

A ranking function is *σ-spam resistant* if, for every graph

Cost of spammer acquiring nodes

$$\frac{\text{Spammer's Cost}}{\text{Spammer's Benefit}} \geq \sigma$$

Contribution: Definition of Spam Resistance

A ranking function is *σ-spam resistant* if, for every graph

Cost of spammer acquiring nodes

$$\frac{\text{Spammer's Cost}}{\text{Spammer's Benefit}} \geq \sigma$$

Sum of rank all nodes owned by spammer

Contribution: Definition of Spam Resistance

A ranking function is *σ-spam resistant* if, for every graph

Cost of spammer acquiring nodes

choice of free nodes,
aquired nodes, and edge changes

$$\min_{\text{spammer choices}} \frac{\text{Spammer's Cost}}{\text{Spammer's Benefit}} \geq \sigma$$

Sum of rank all nodes owned by spammer

Contribution: Definition of Spam Resistance

A ranking function is σ*-spam resistant* if, for every graph

Cost of spammer acquiring nodes

choice of free nodes,
aquired nodes, and edge changes

$$\max_{\text{choice of cost function}} \; \min_{\text{spammer choices}} \frac{\text{Spammer's Cost}}{\text{Spammer's Benefit}} \geq \sigma$$

Sum of rank all nodes owned by spammer

Contribution: Definition of Spam Resistance

A ranking function is σ-*spam resistant* if

$$\max_{\text{choice of cost function}} \; \min_{\text{spammer choices}} \frac{\text{Spammer's Cost}}{\text{Spammer's Benefit}} \geq \sigma$$

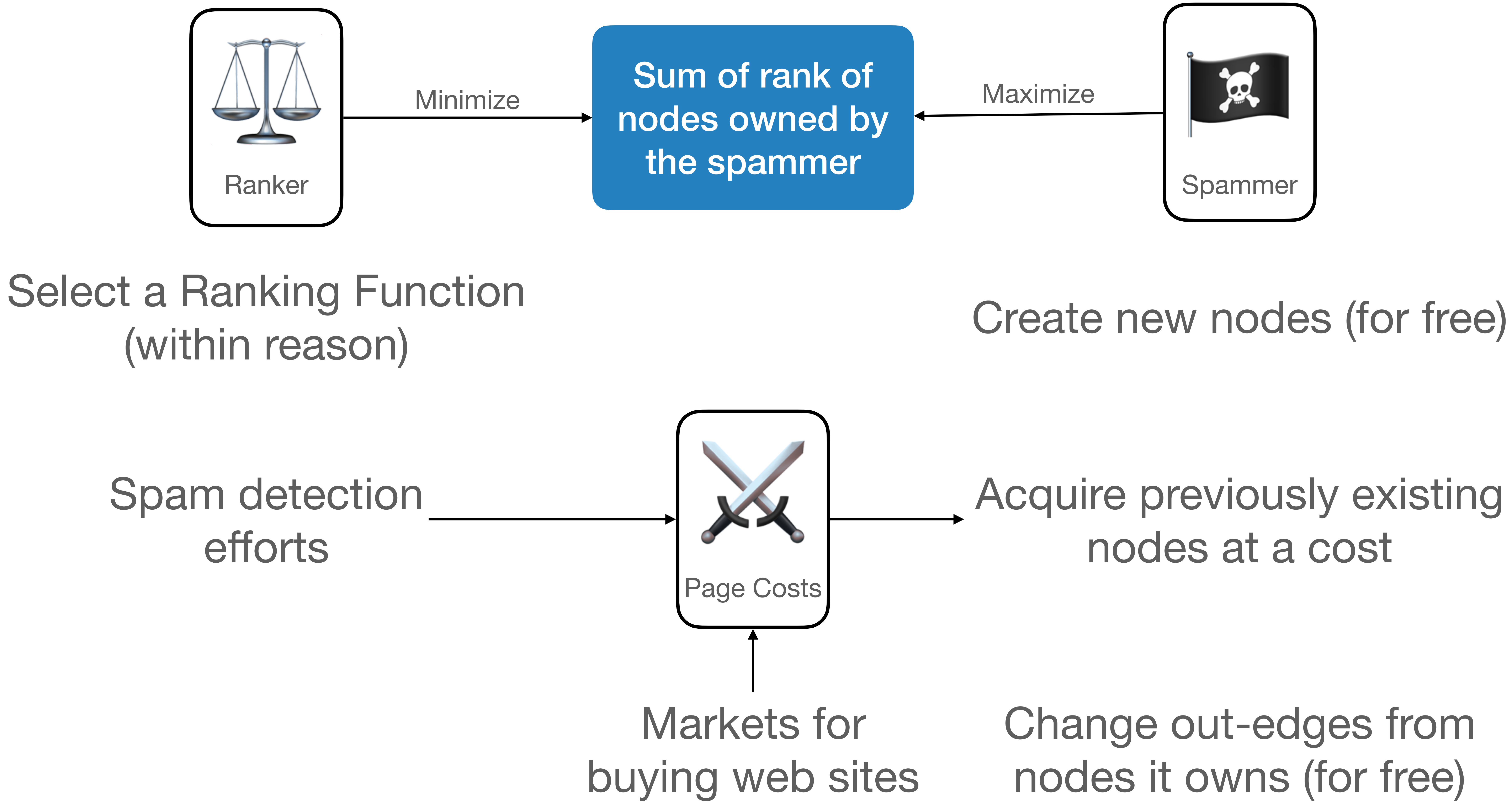
Lemma: UPR is 0-spam resistant

Theorem: PPR with reset probability ε is ε-spam resistant

- Conjecture: This is the best any PageRank can do over all choices of reset vectors

So why not use PPR?

What can the Ranker and Spammer do?



What can the Ranker and Spammer do?

Ranker

Minimize

Sum of rank of nodes owned by the spammer

Maximize

Spammer

Select a Ranking Function
(within reason)

Create new nodes (for free)

Spam detection efforts

Page Costs

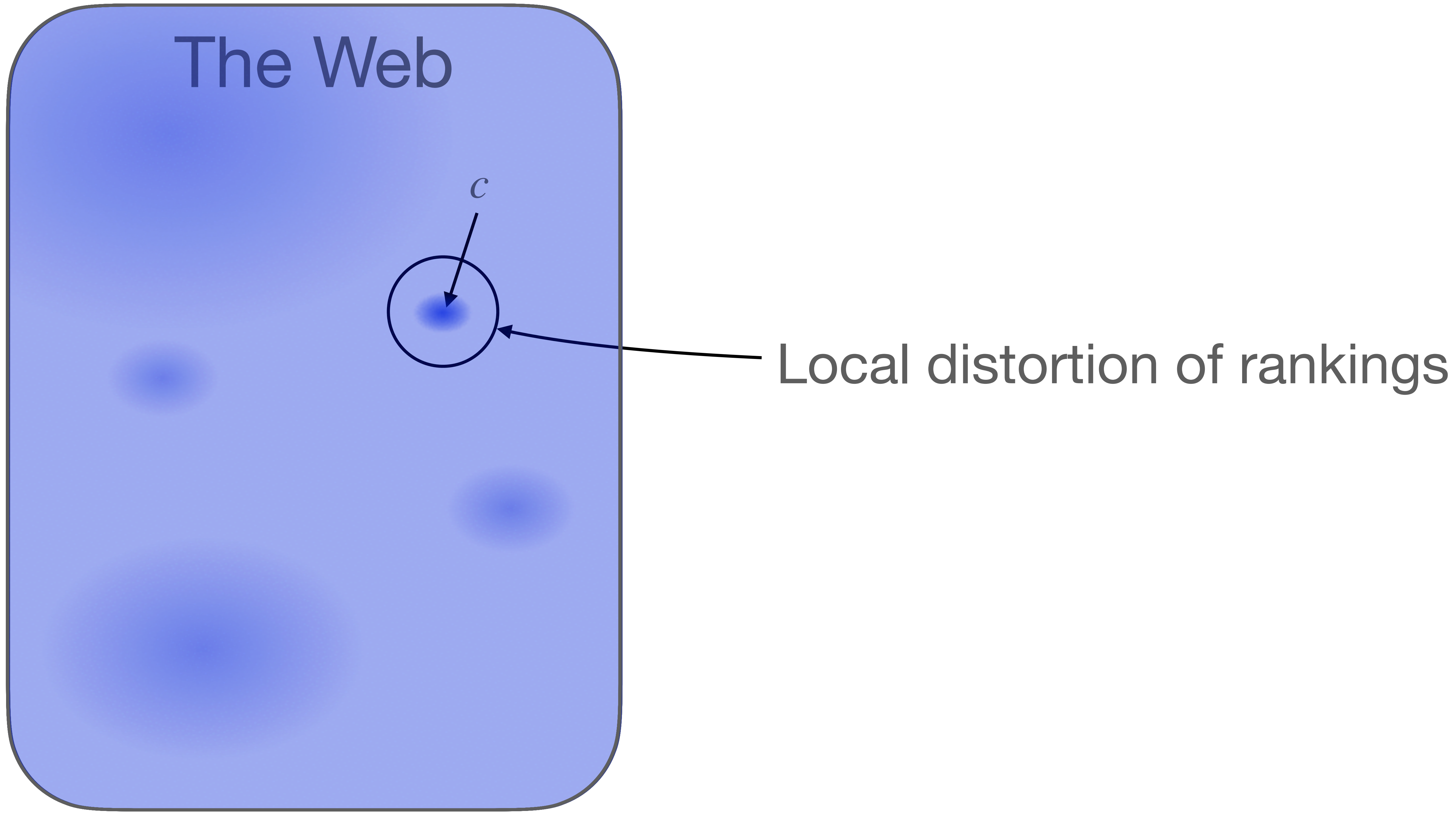
Acquire previously existing nodes at a cost

Markets for buying web sites

Change out-edges from nodes it owns (for free)

Personalized PageRank has local distortion

The nodes just downstream of c get very high rank



PPR_c

Contribution: Defining distortion of a ranking function

Following Brin & Page:

- On ergodic graphs, stationary distribution is the **reference** rank

We define *distortion* as the max multiplicative error vs reference rank

$$\max_{\text{ergodic graphs } G=(V,E)} \quad \max_{v \in V} \quad \max \left\{ \overbrace{\frac{\text{rank of } v}{\text{reference rank of } v}}^{\text{stretch}}, \overbrace{\frac{\text{reference rank of } v}{\text{rank of } v}}^{\text{contraction}} \right\}^{*}$$

*for technical reasons, we round up all ranks to something reasonable like $1/n^{O(1)}$

Partial results

Theorem: UPR has poor spam resistance and poor distortion

Theorem: PPR has good spam resistance and poor distortion

Partial results

Theorem: UPR has poor spam resistance and poor distortion

Theorem: PPR has good spam resistance and poor distortion

That’s progress, but still need to deal with distortion

Partial results

Theorem: UPR has poor spam resistance and poor distortion

Theorem: PPR has good spam resistance and poor distortion

That’s progress, but still need to deal with distortion

What does a grad student* do when their idea doesn’t work?

Partial results

Theorem: UPR has poor spam resistance and poor distortion

Theorem: PPR has good spam resistance and poor distortion

That’s progress, but still need to deal with distortion

What does a grad student* do when their idea doesn’t work?

Come up with a workaround!

Partial results

Theorem: UPR has poor spam resistance and poor distortion

Theorem: PPR has good spam resistance and poor distortion

That's progress, but still need to deal with distortion

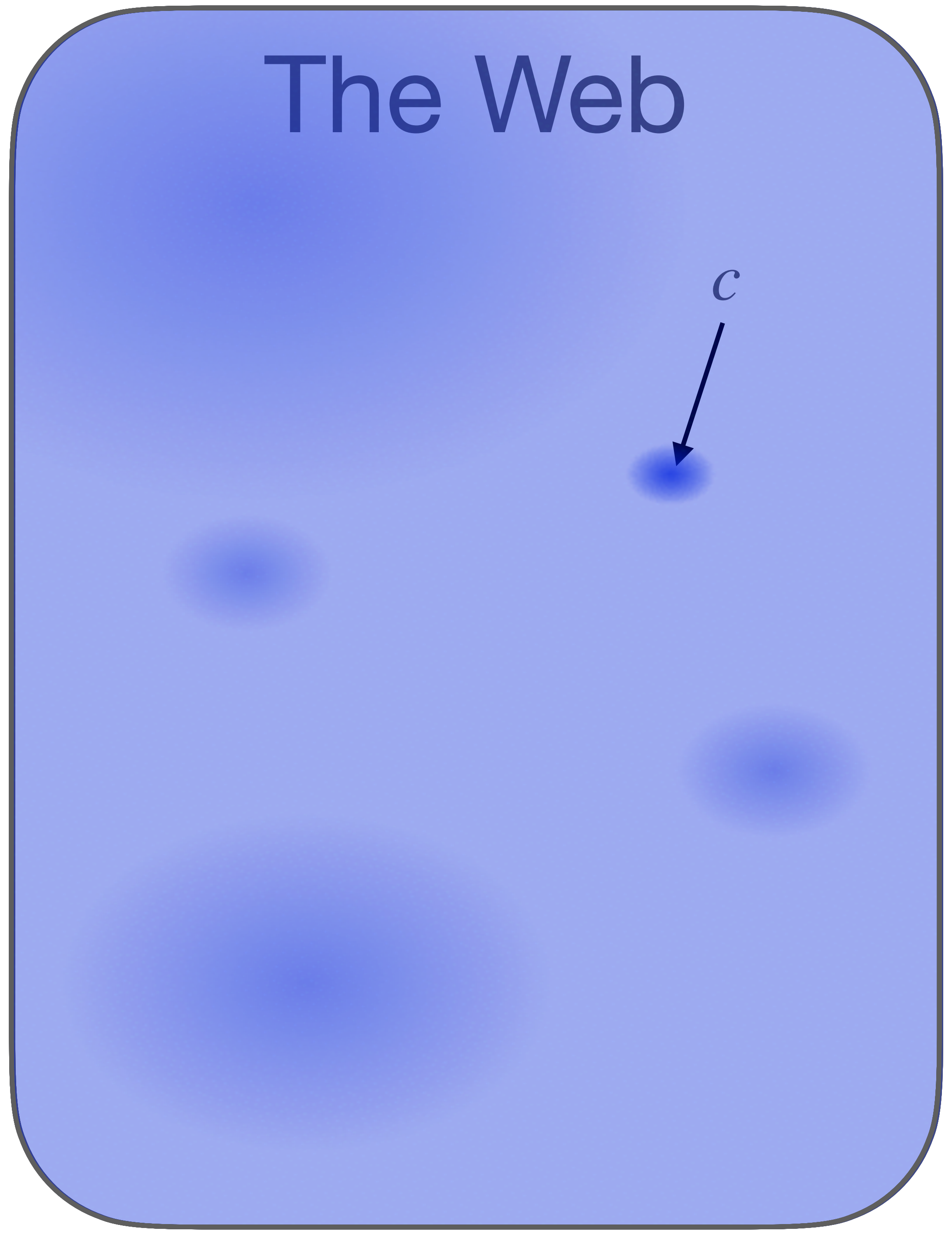
What does a grad student* do when their idea doesn't work?

Come up with a workaround!

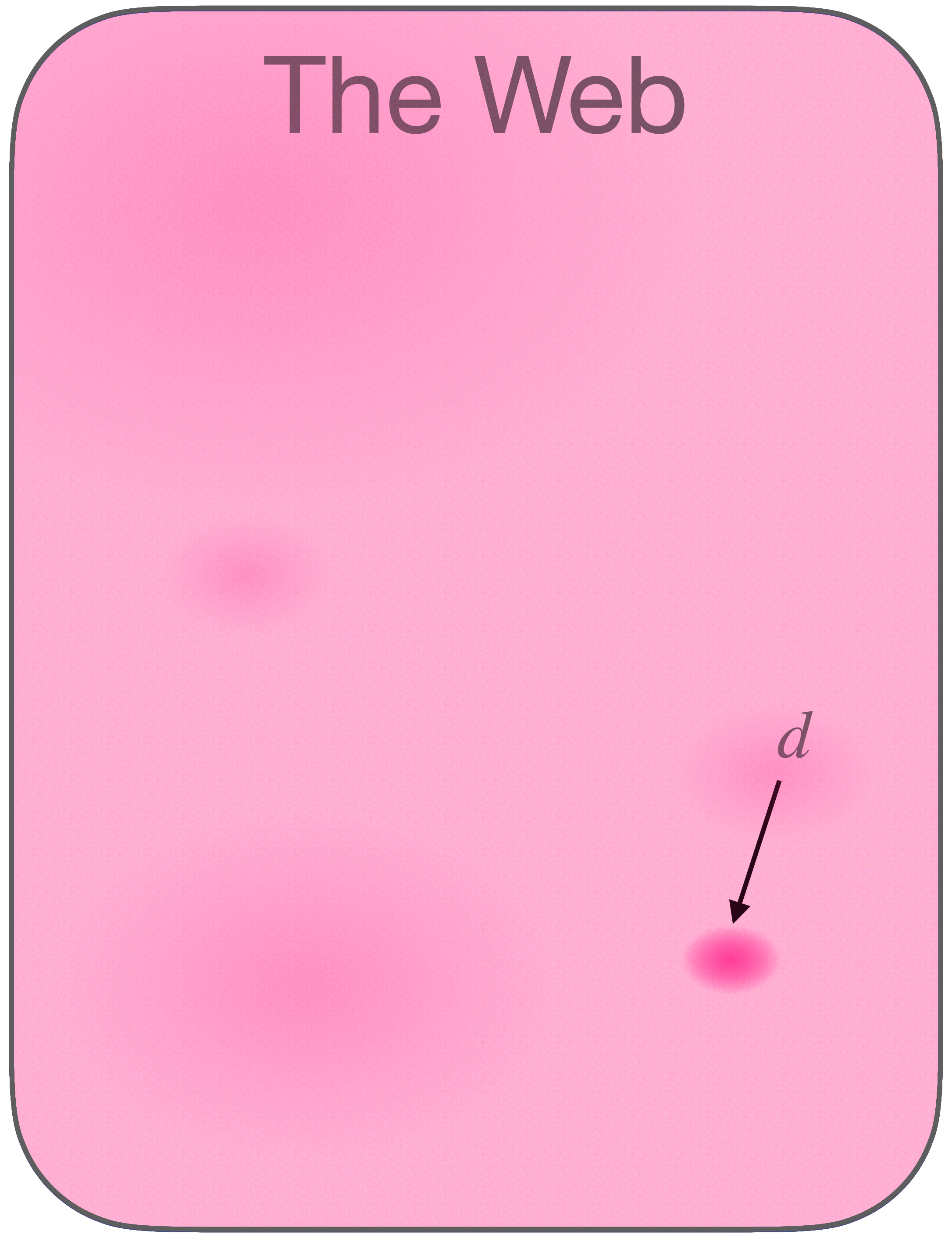
*Again, Brin and Page!

Workaround to fix distortion of PPR

Compute two PPRs: one with center c and one with center d

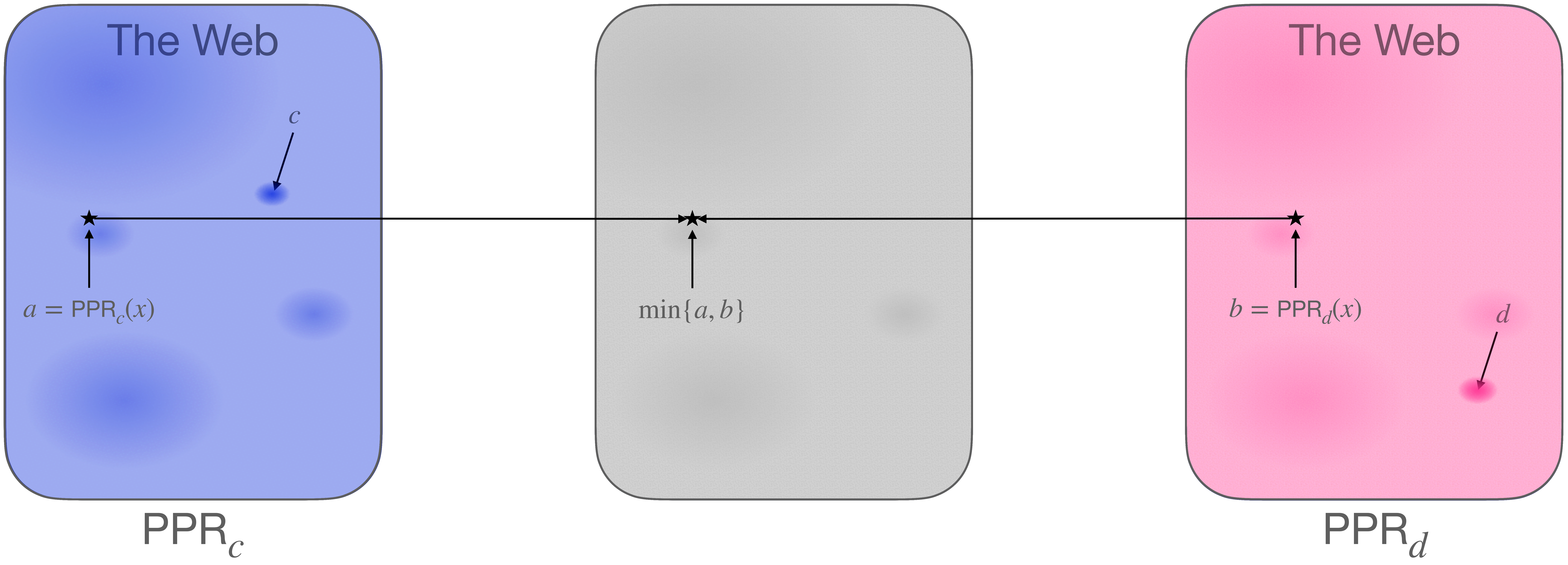


PPR_c



PPR_d

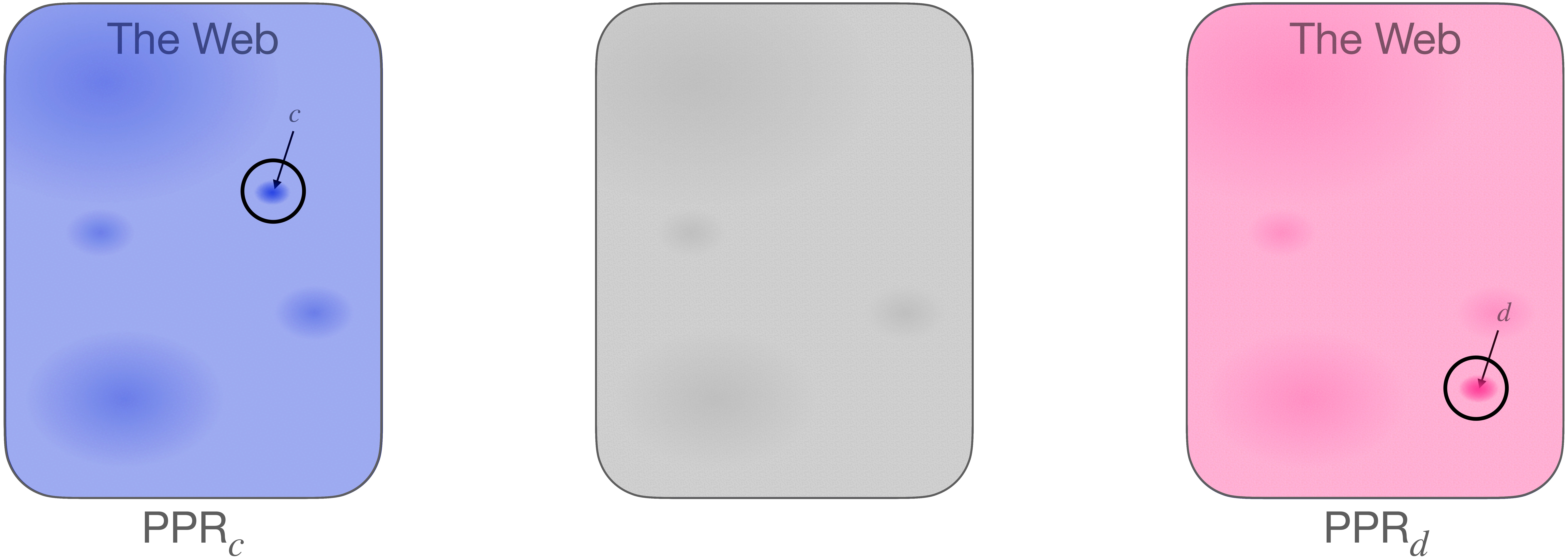
Compute two PPRs: one with center c and one with center d



Compute the (normalized) component-wise min!

Reducing distortion of PPR

Compute two PPRs: one with center c and one with center d

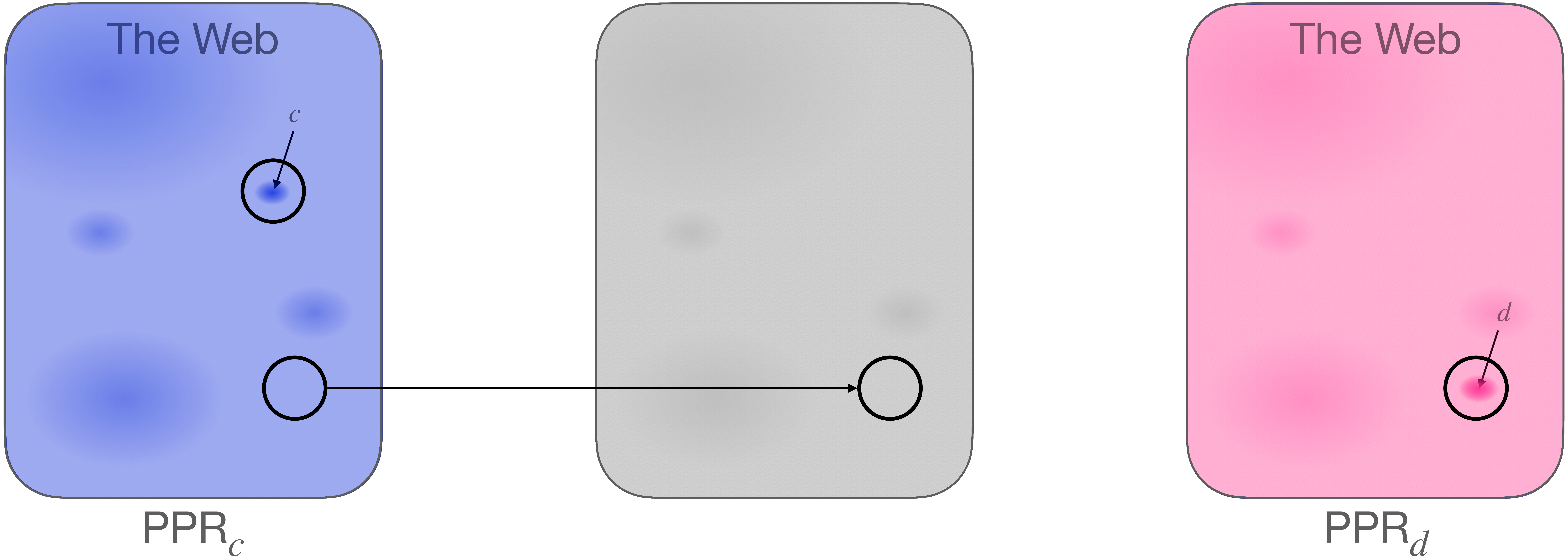


Compute the component-wise min!

- PPR_c will kill distortion around d and vice versa

Reducing distortion of PPR

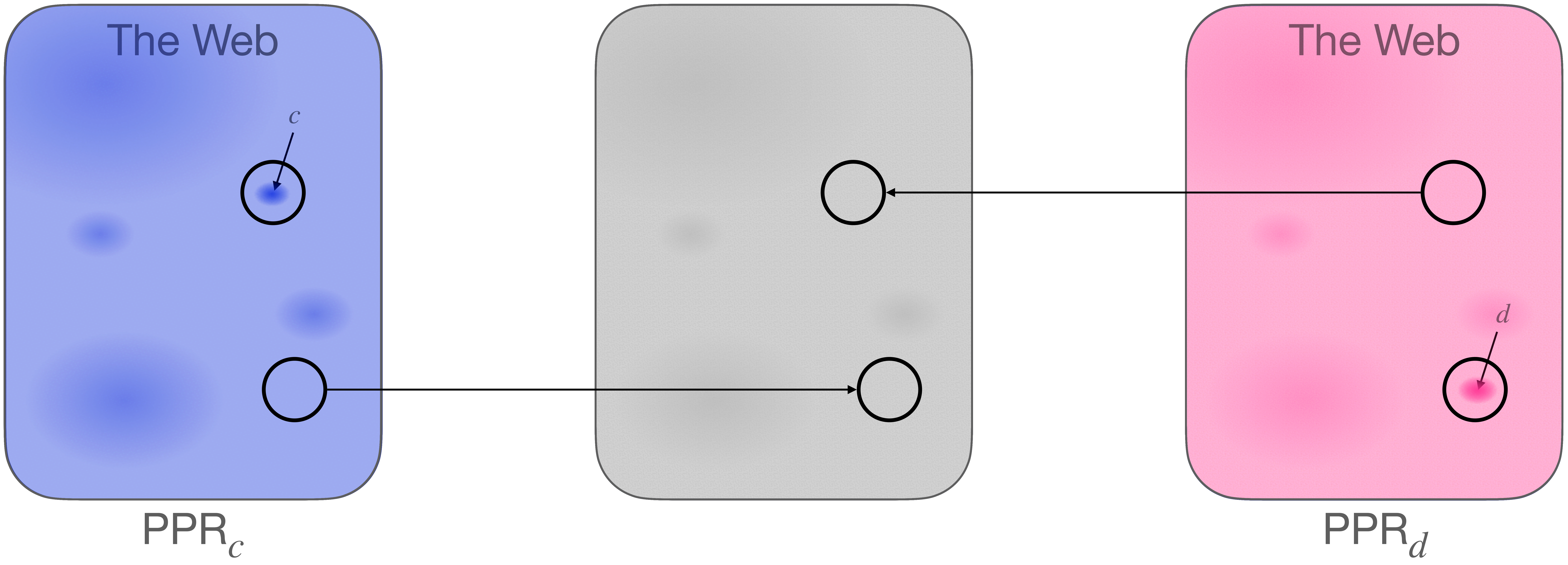
Compute two PPRs: one with center c and one with center d



Compute the component-wise min!

- PPR_c will kill distortion around d and vice versa

Compute two PPRs: one with center c and one with center d



Compute the component-wise min!

- PPR_c will kill distortion around d and vice versa

And now for main theorems

Main theorems

Main theorems

Theorem: For almost all choices of centers, Min-PPR has low distortion!

- Almost all = from a sensible distribution (see paper)

Main theorems

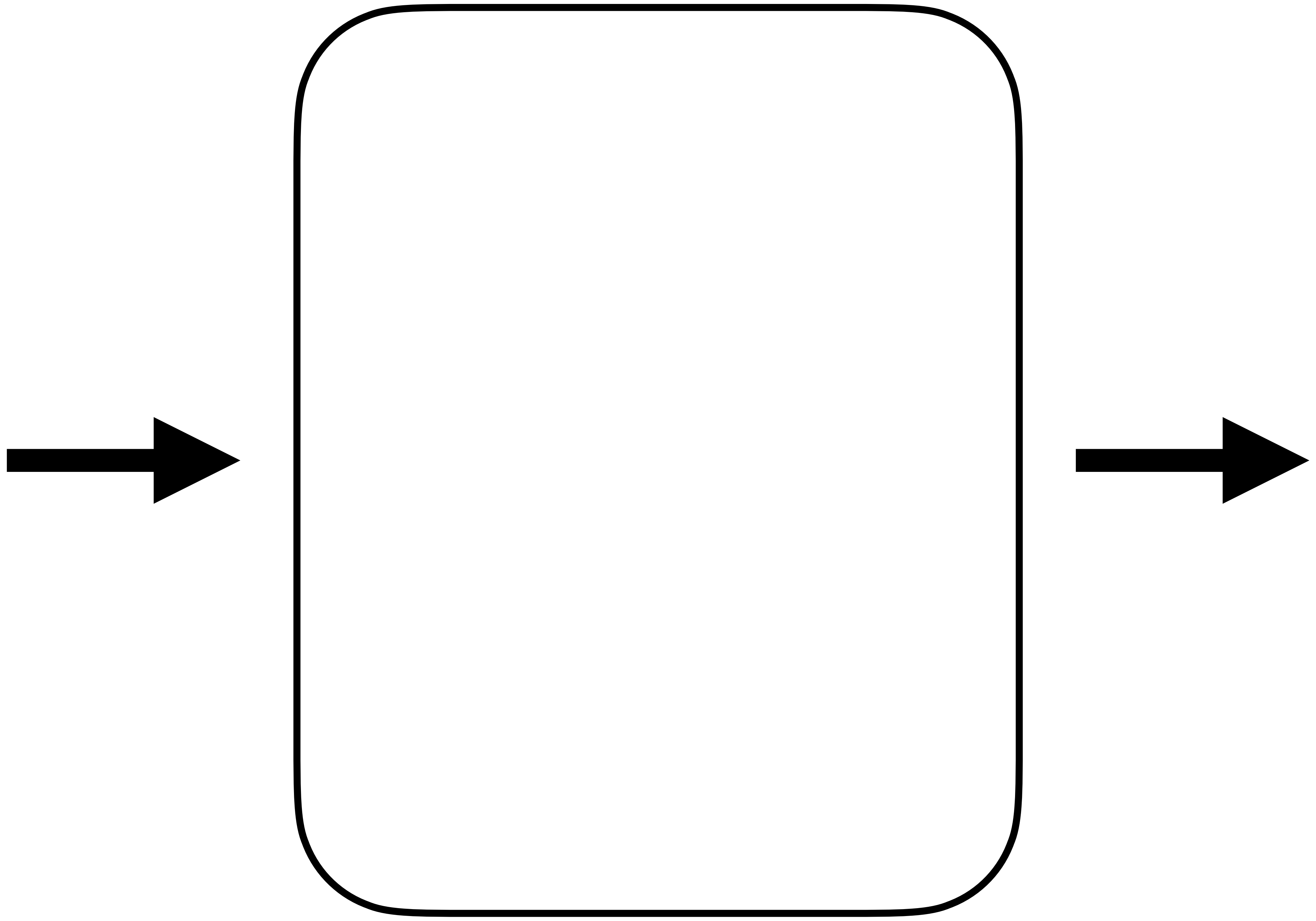
Theorem: For almost all choices of centers, Min-PPR has low distortion!

- Almost all = from a sensible distribution (see paper)

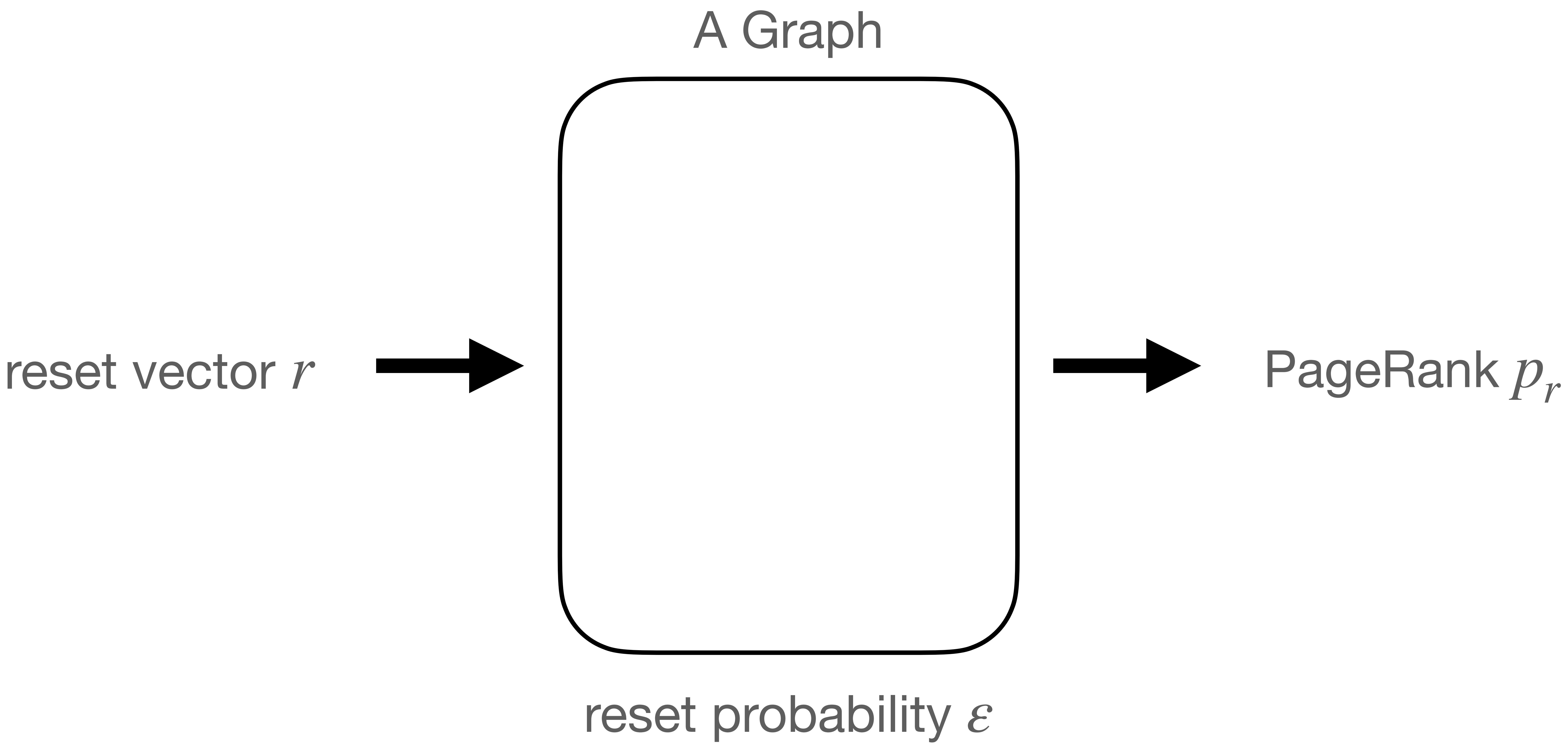
Theorem: Min-PPR with k centers is ε/k-spam resistant!

- So if k is small, Min-PPR is almost as spam resistant as PPR
- And the proof gives a concrete (natural) cost function
- The cost function tells the rankers where to spend their spam-detection efforts

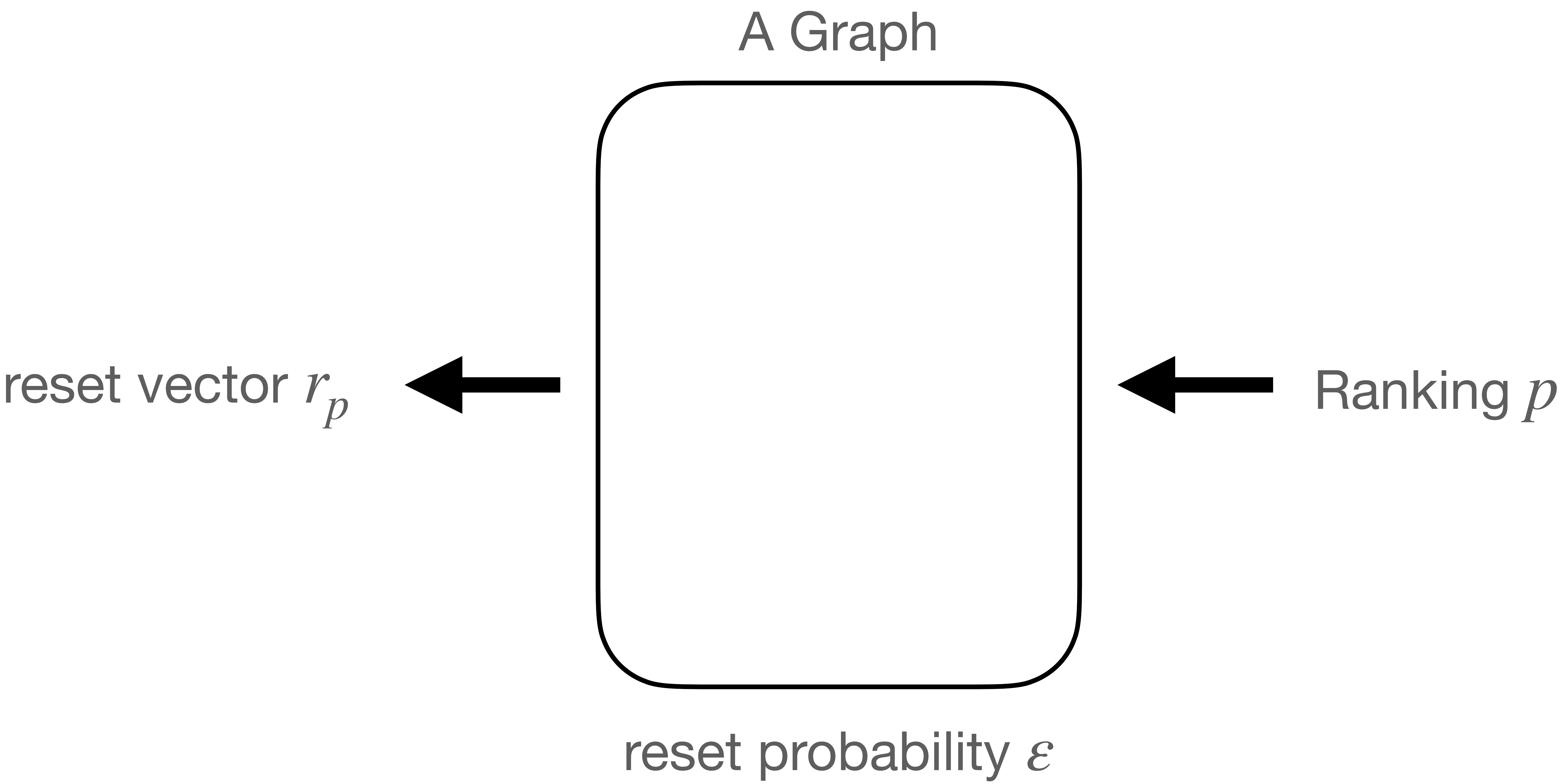
One last thing


A Graph
reset vector r
PageRank p_r
reset probability ε

When is a ranking a PageRank?

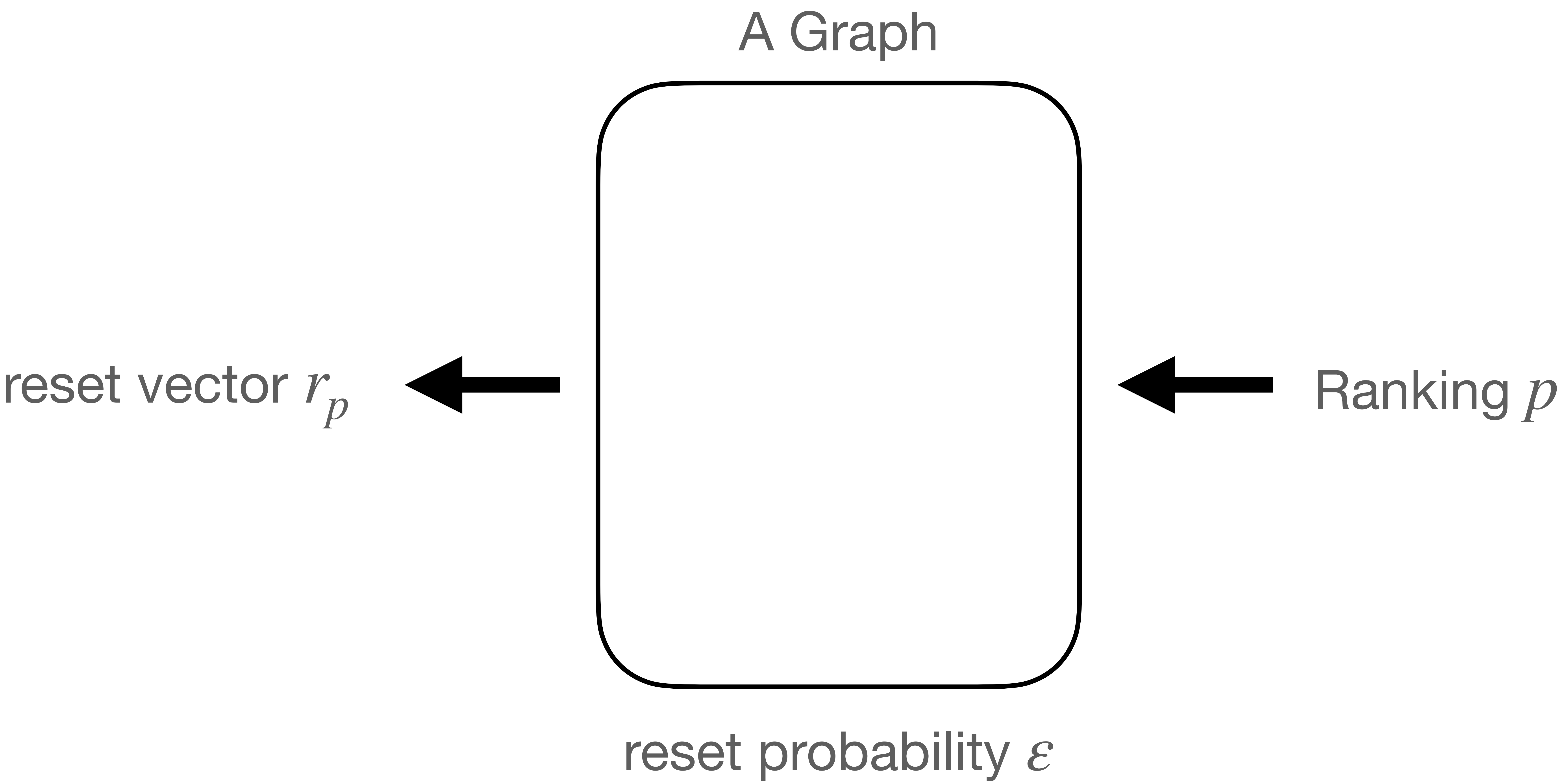


Technical Nugget



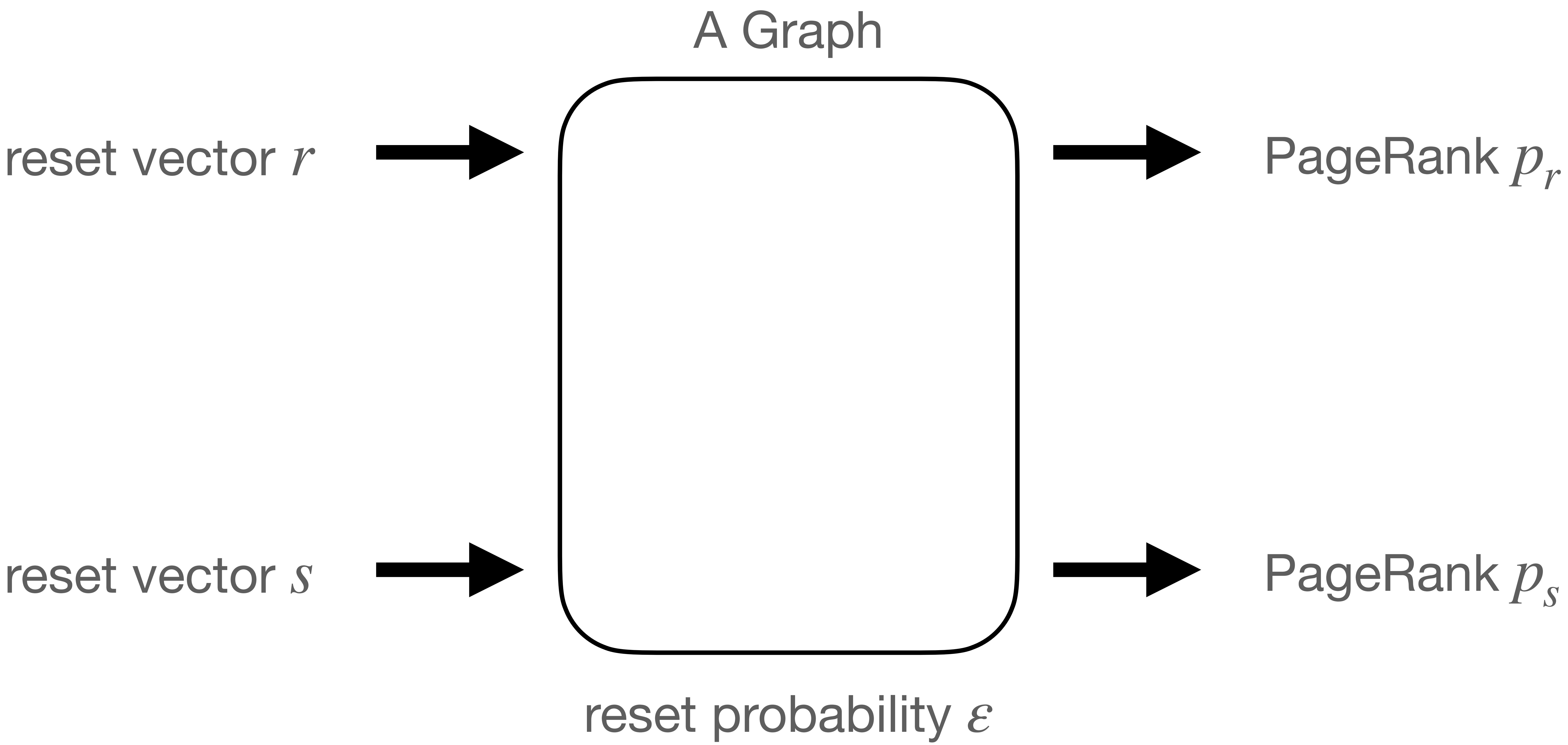
Technical Nugget

If we fix ε, then we can invert from p to r_p

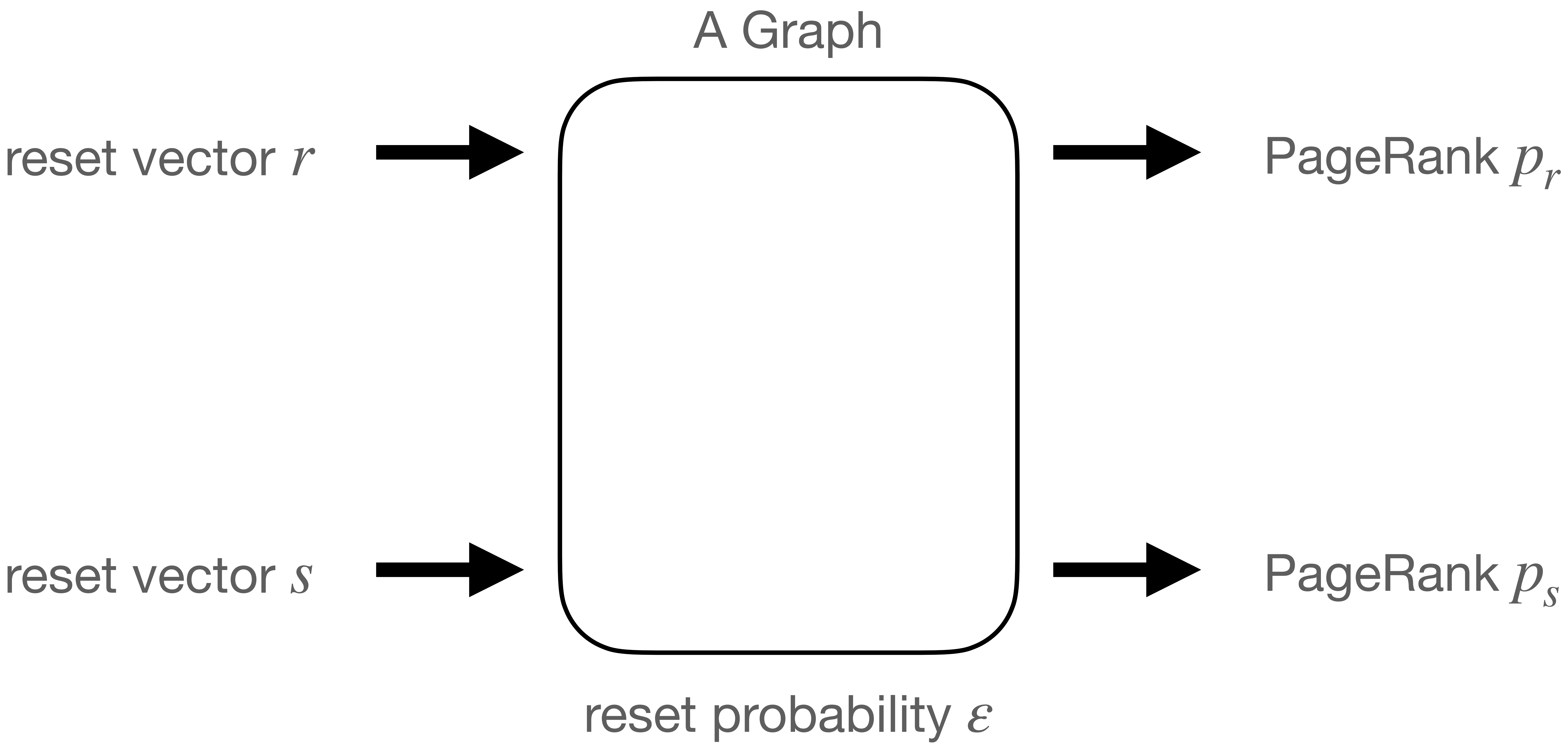


Theorem: p is a PageRank iff the computed r_p has no negative values.

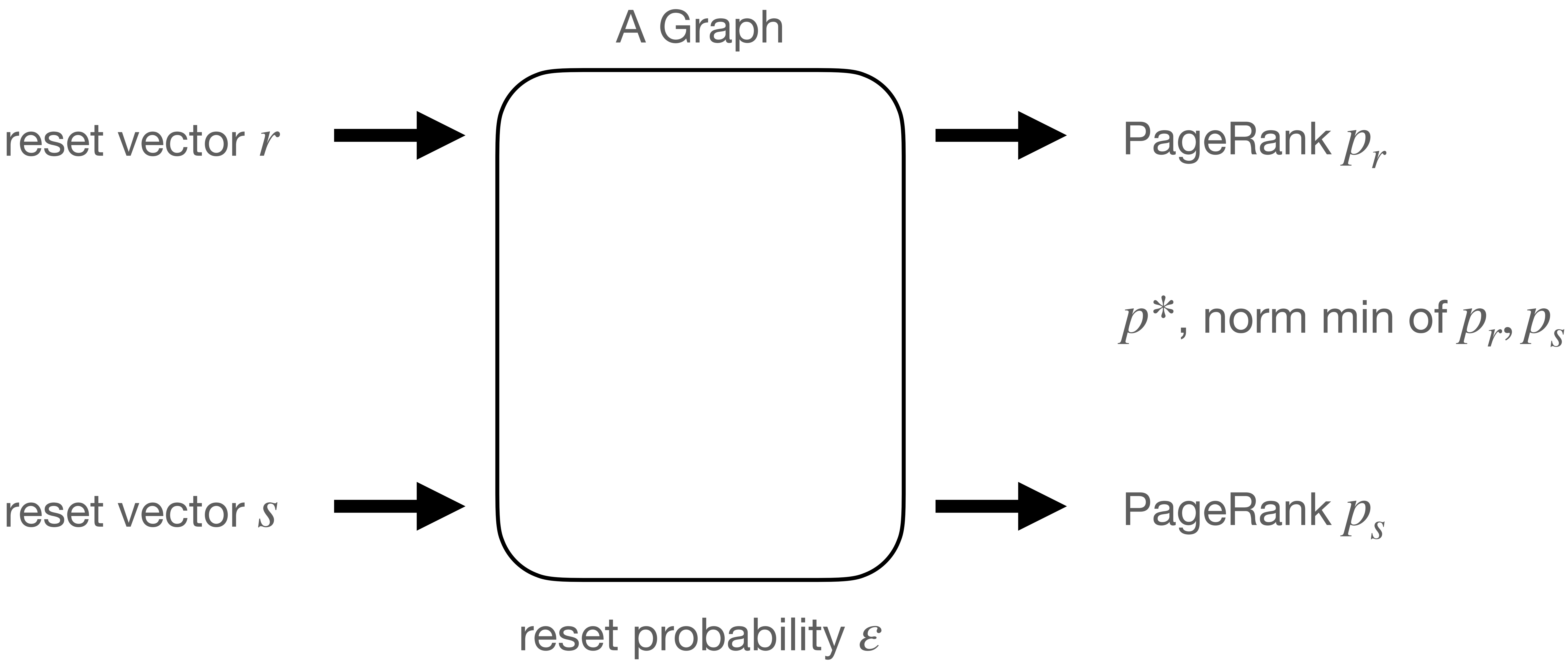
Technical Nugget



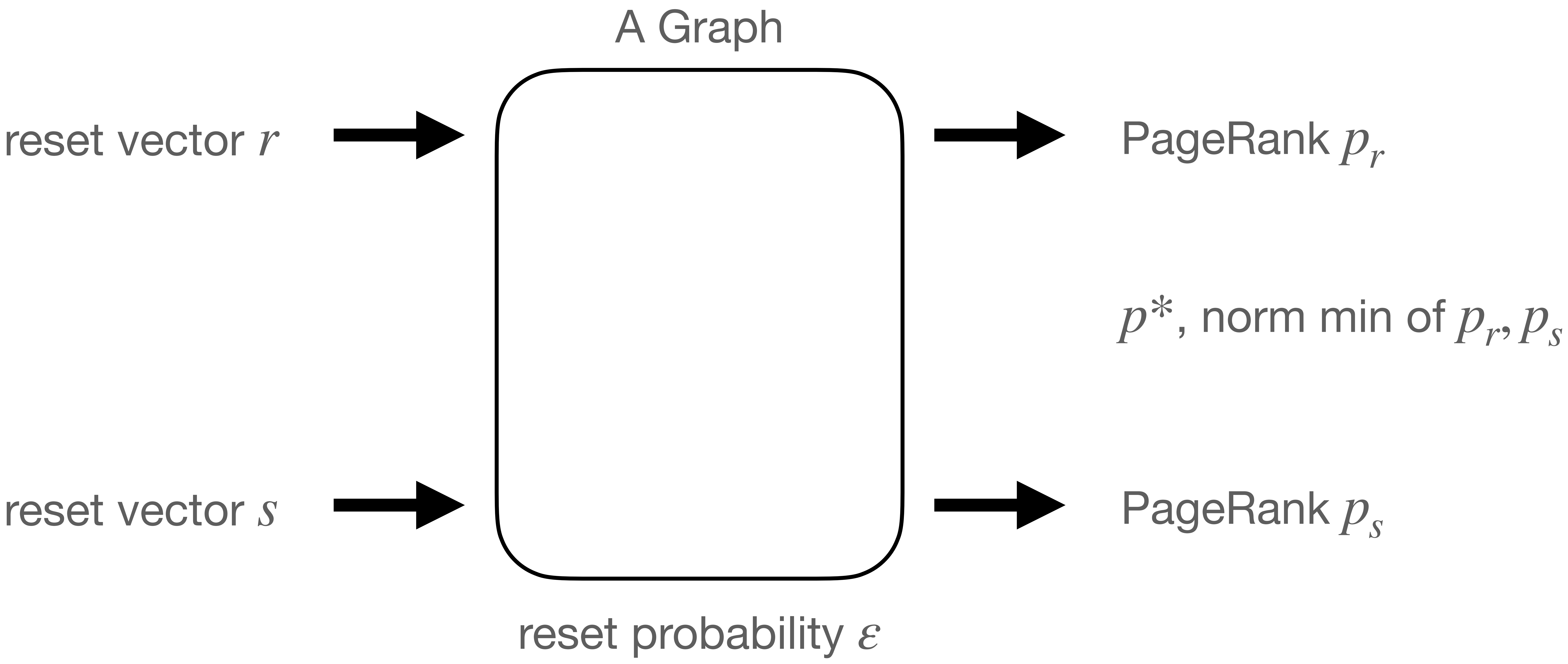
Fun fact about PageRank and component-wise min



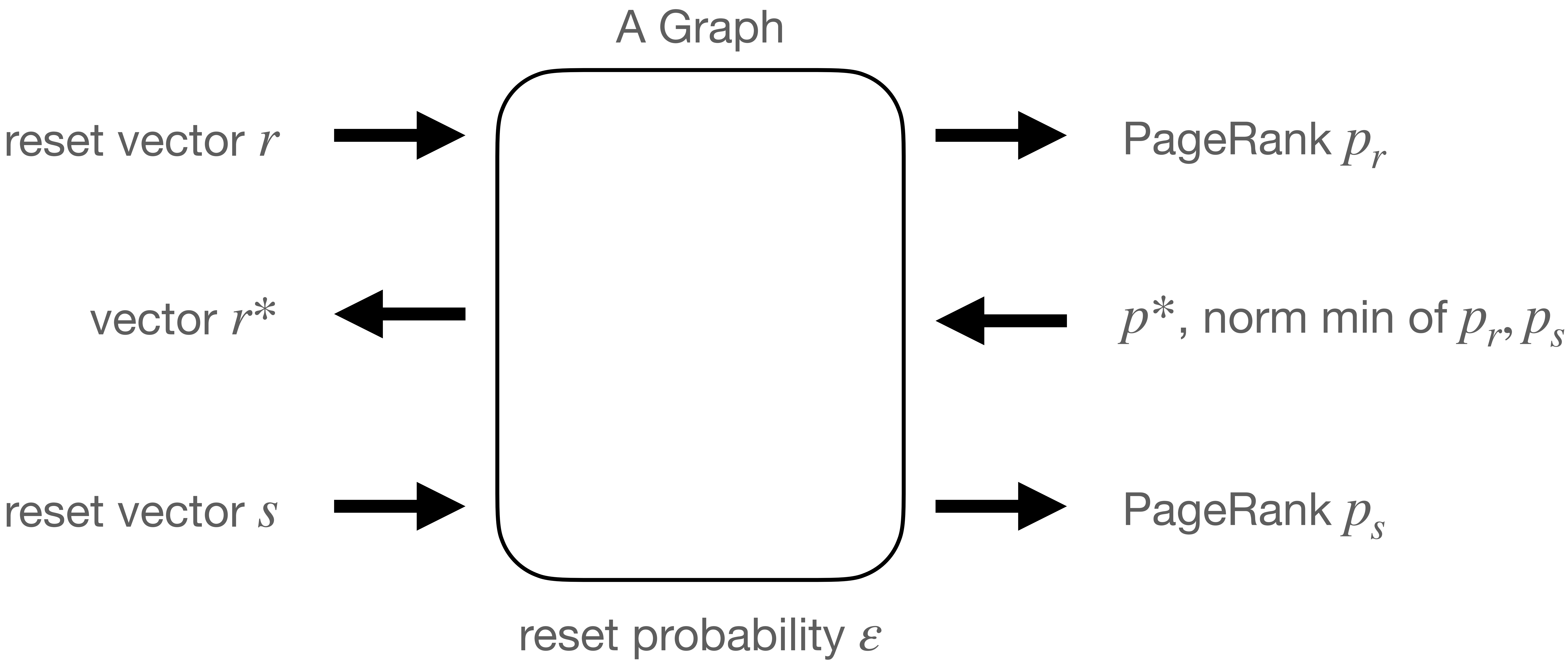
Technical Nugget



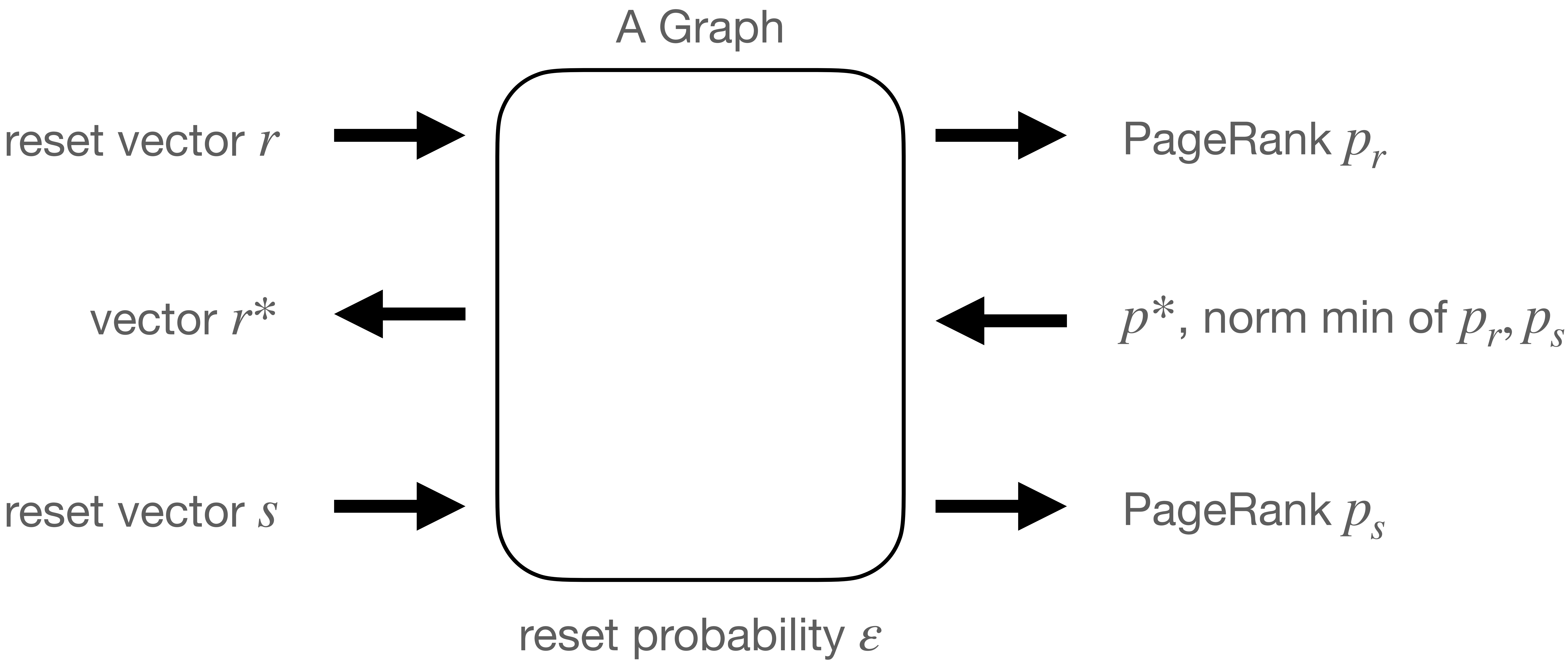
Fun fact about PageRank and component-wise min



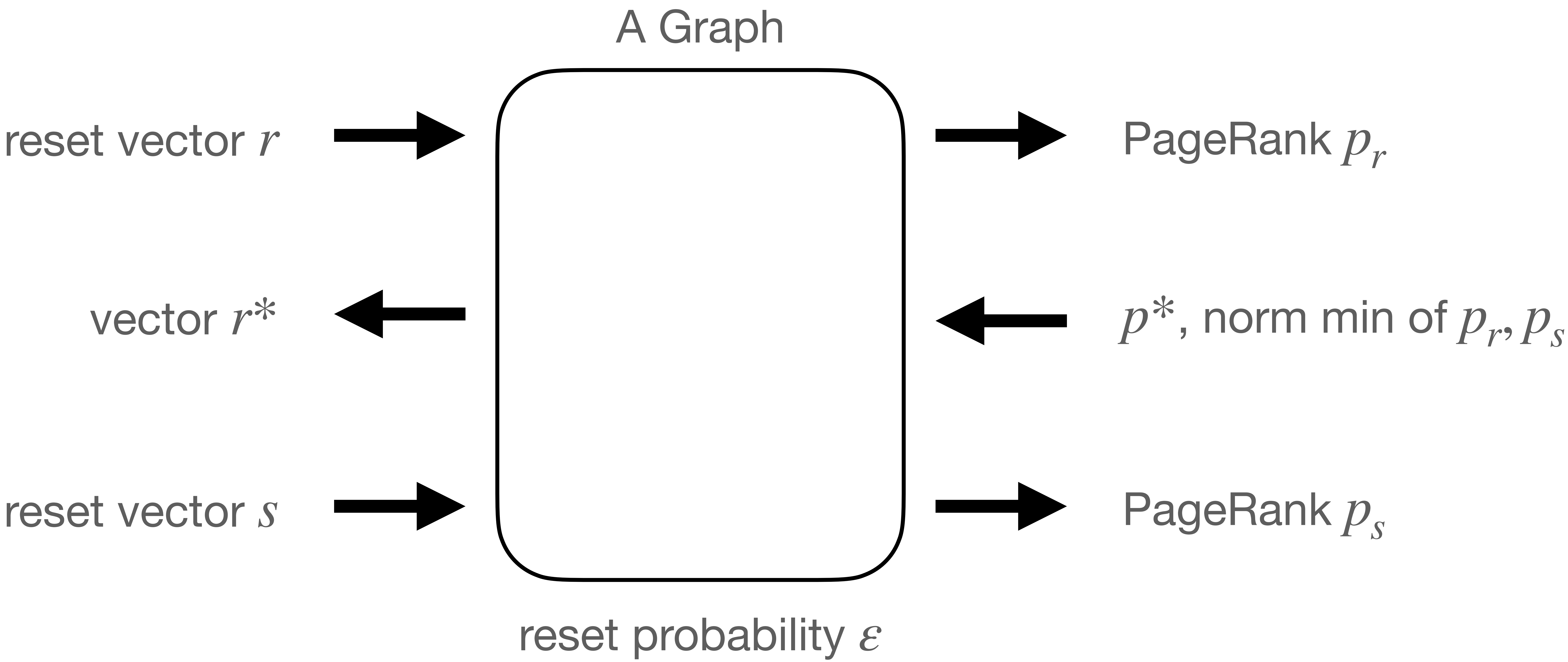
Technical Nugget



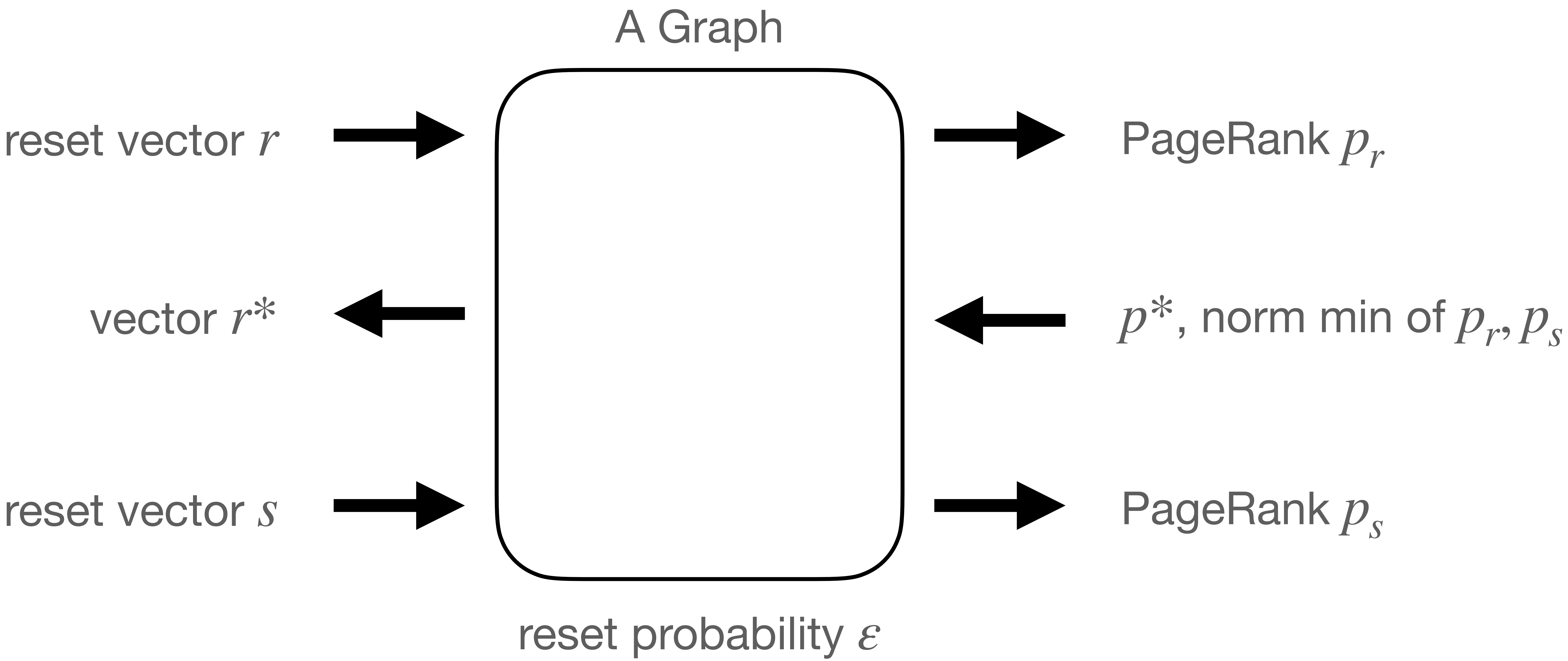
Fun fact about PageRank and component-wise min



Technical Nugget

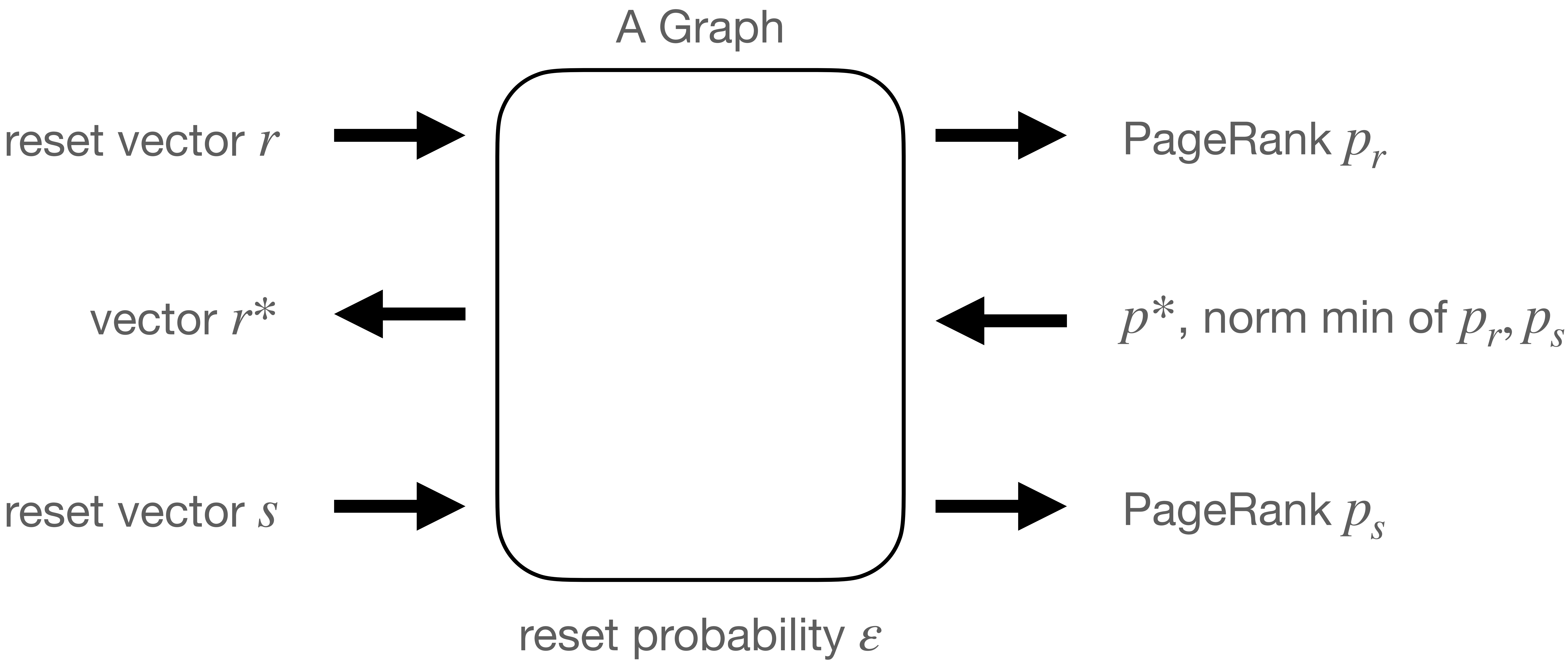


Fun fact about PageRank and component-wise min

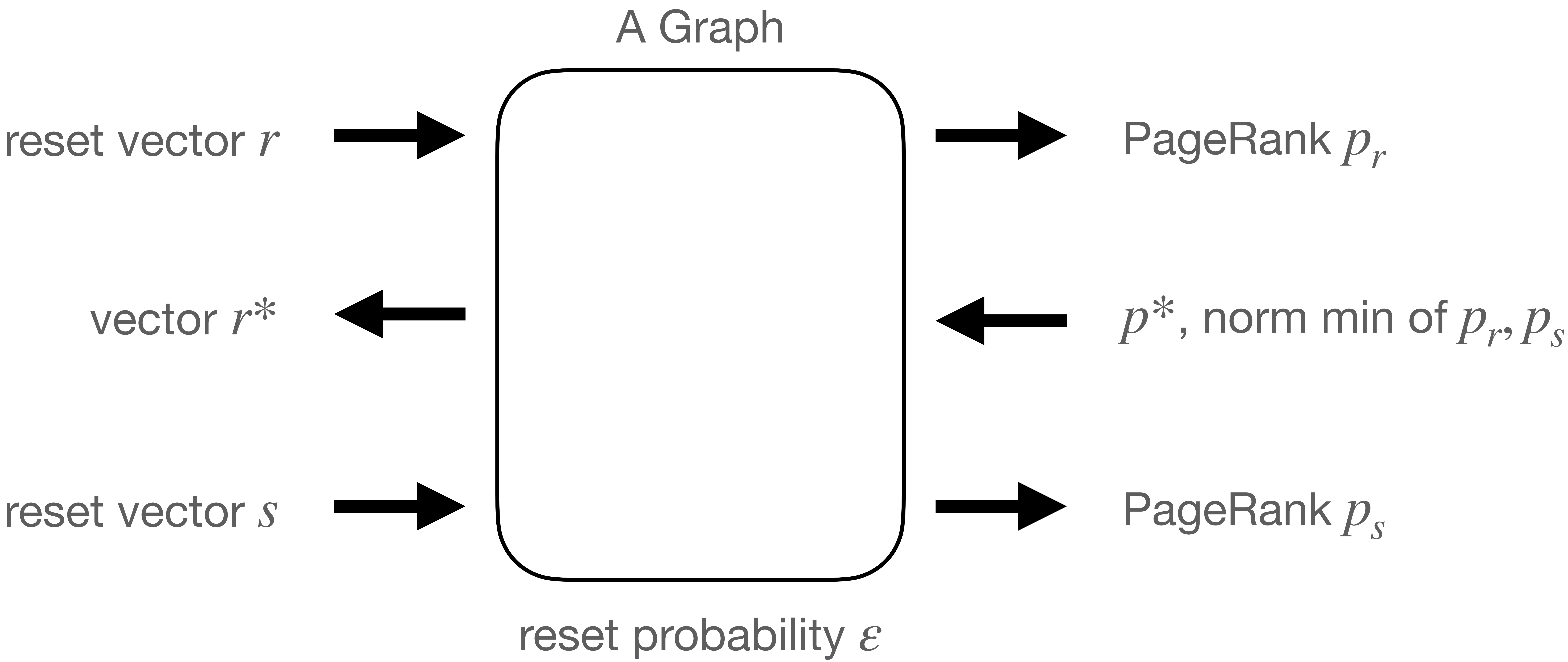


Theorem: p^* is a PageRank (or all 0) (eqv., r^* is non-negative)

Technical Nugget



Fun fact about PageRank and component-wise min



Theorem: PageRank vectors is closed under norm. component-wise min

Summary

Summary

We formalize web ranker-spammer interaction as a game

Summary

We formalize web ranker-spammer interaction as a game

We observe that all ranking functions in the literature have poor distortion or poor spam resistance (or both)

Summary

We formalize web ranker-spammer interaction as a game

We observe that all ranking functions in the literature have poor distortion or poor spam resistance (or both)

We prove that the heuristic used by Google has good resistance and good distortion

- Technical nugget: we use the closure property of PageRank extensively in our proofs

Further questions

More work to be done on the web ranking spamming game

- Formalize setting the cost function?

Spamming is ubiquitous. Can we analyze games for:

- Wikipedia spamming?
- Citation spamming? H-index spamming?

Further questions

More work to be done on the web ranking spamming game

- Formalize setting the cost function?

Spamming is ubiquitous. Can we analyze games for:

- Wikipedia spamming?
- Citation spamming? H-index spamming?

Goodhart's Law

When a measure becomes a target, it ceases to be a good measure.

Further questions

More work to be done on the web ranking spamming game

- Formalize setting the cost function?

Spamming is ubiquitous. Can we analyze games for:

- Wikipedia spamming?
- Citation spamming? H-index spamming?

Goodhart's Law

When a measure becomes a target, it ceases to be a good measure.

(unless you can prove spam resistance)