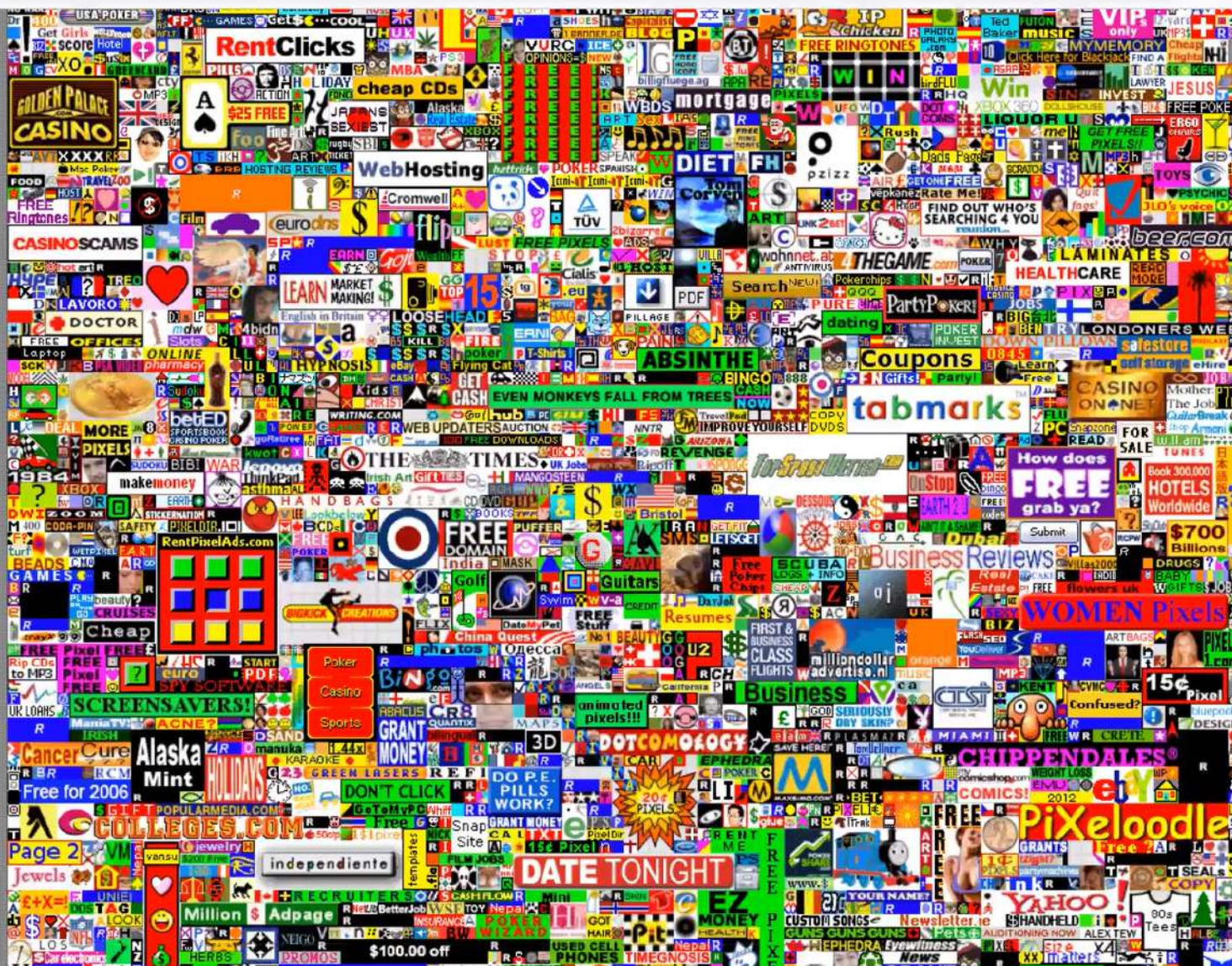


Related aspect...

◆ "Marketing 2.0"

◆ 1.000.000 ...





Compare...

◆*pod





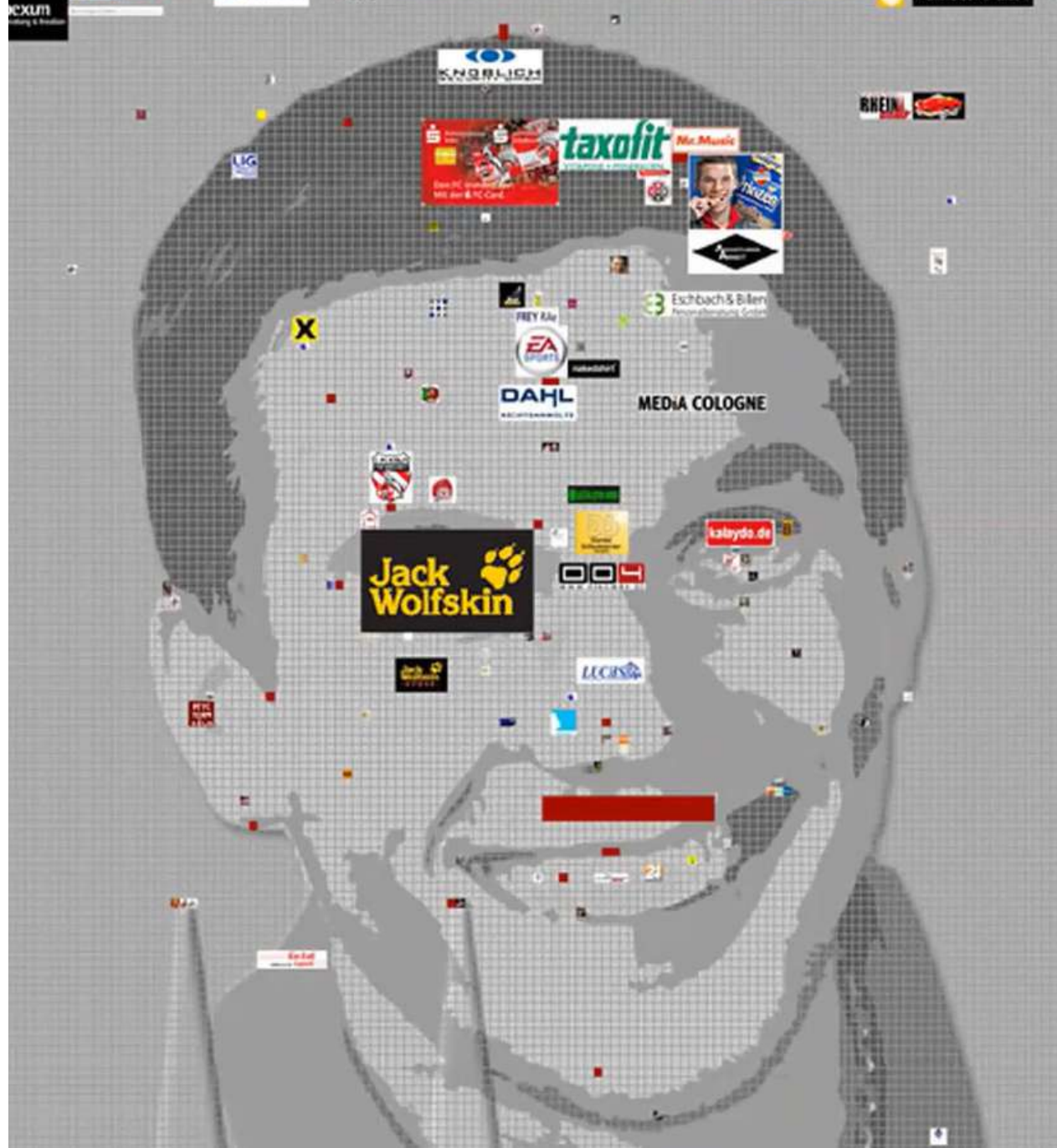
Lukas Podolski

Traumtor zum 3:0
Köln-Saarbrücken

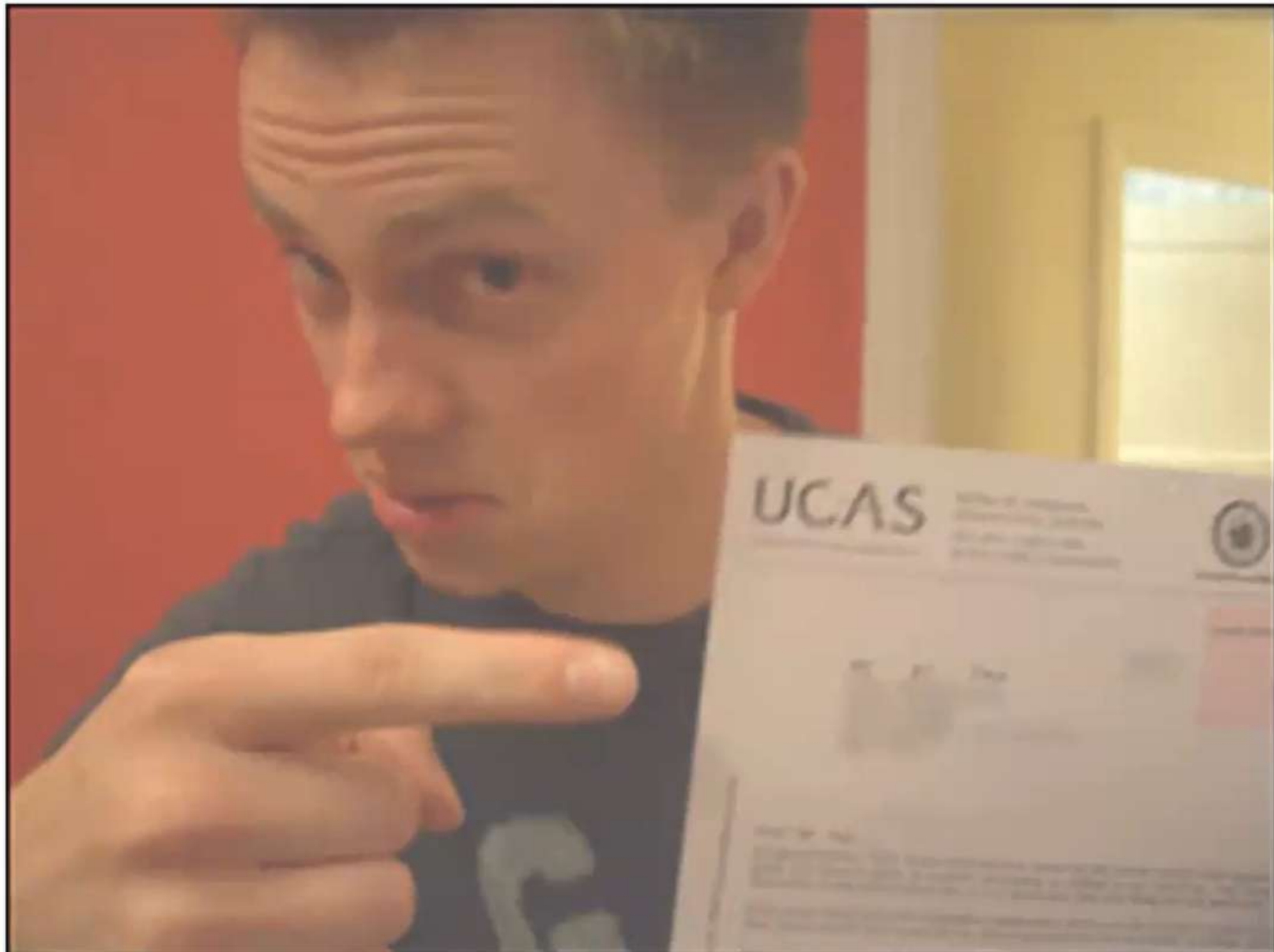
D:SF

Kreis II - 1. Daxsteinen 38:30





Alex Tew, “genius by chance”...!



Compare...

◆ *pod

◆ And so??



Let's analyze...

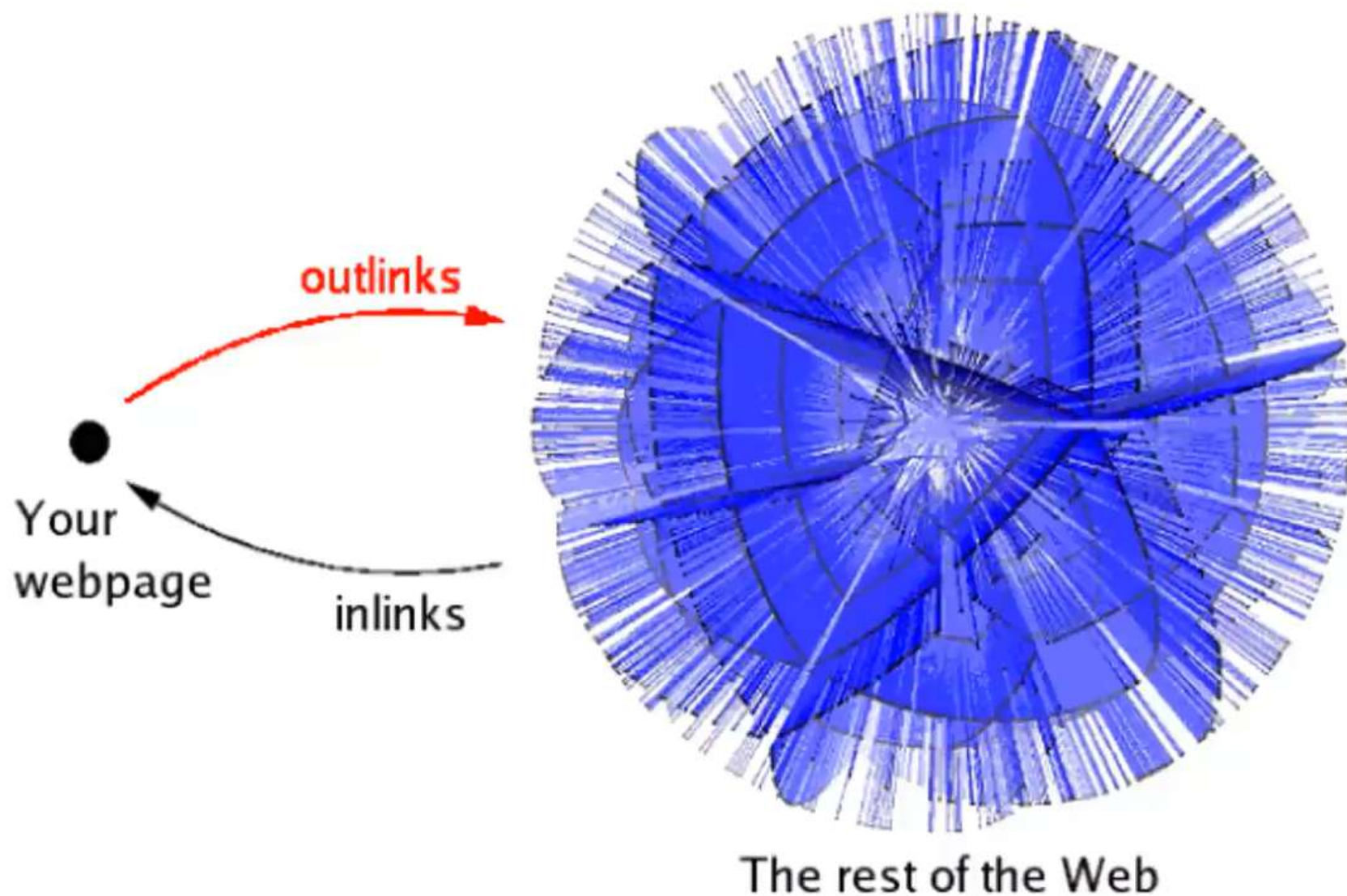
- ◆ ... the market success of the million dollar page



Let's now see the other direction of links



Strategic move: Adding outlinks



What happens in this case?

- ◆ Just reason by symmetry
- ◆ The pagerank gets lower (\leq) because some flow goes out, we cannot augment flow with an outgoing channel



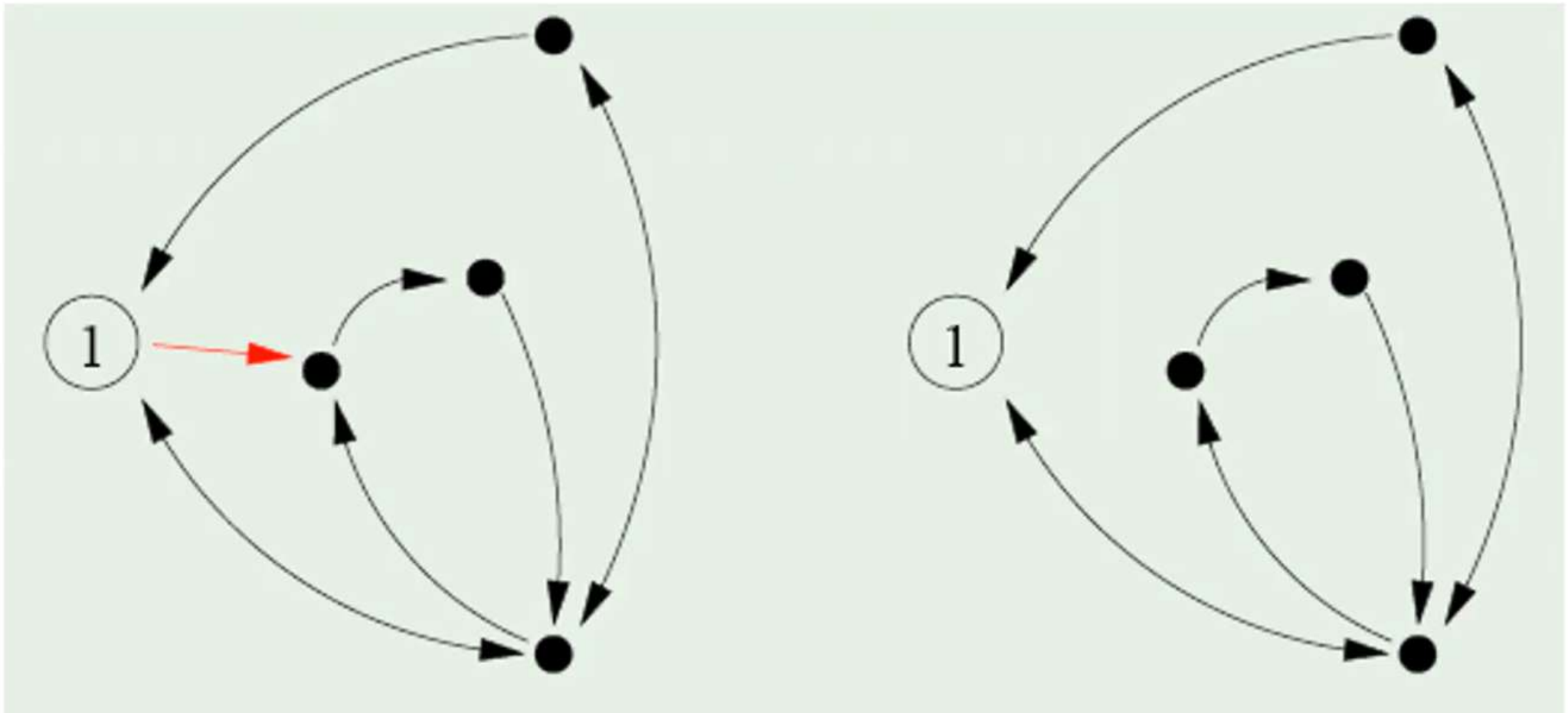
Property: solidity



- ◆ Property of ranking measure, to be “***solid***”: adding outlinks to a page doesn’t cause an increase of the hypertextual score
- ◆ → The spamdex effort cannot just be ***local***

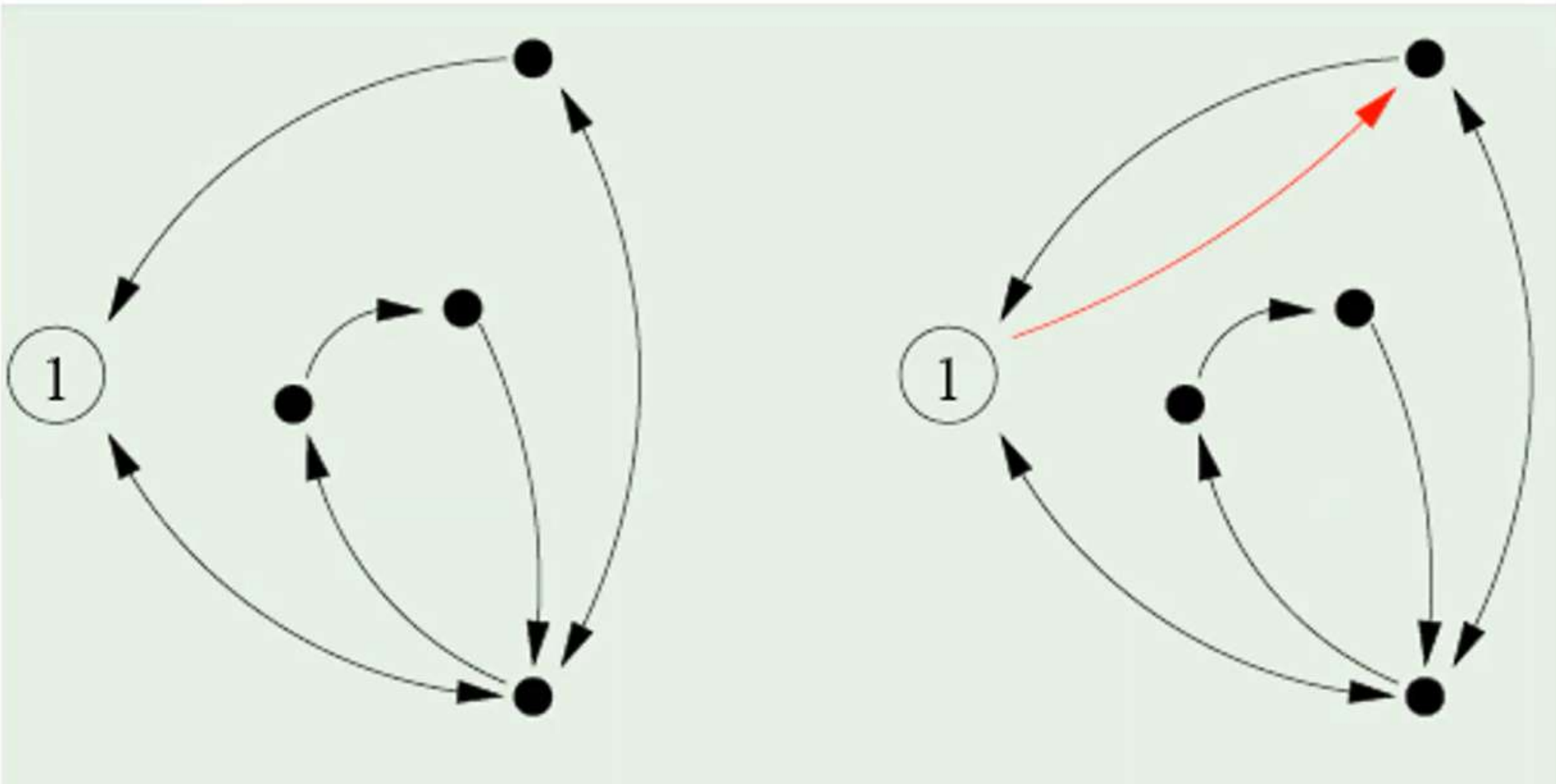
Example

- ◆ **0.196** initially, passing to
- ◆ **0.182**

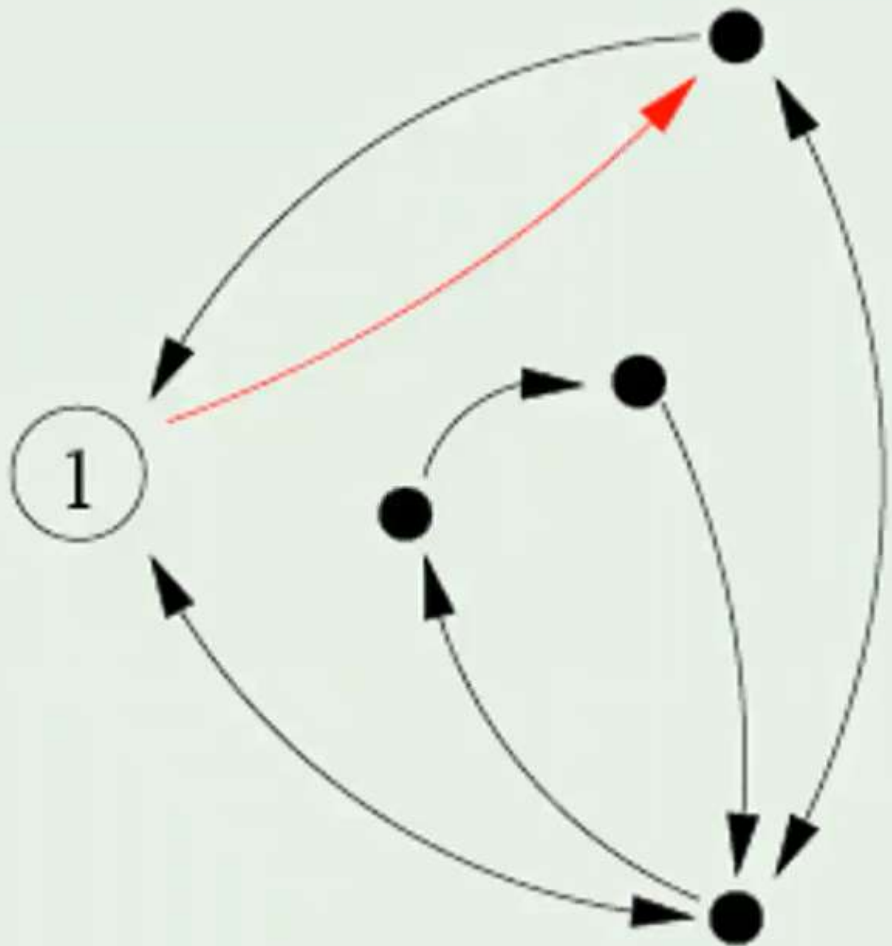


Other example

- ◆ **0.196** initially, passing to
- ◆ **0.211**



Response	Percentage
Yes, the current government is responsible	85%
No, the crisis is not the government's fault	15%



Moral

- ◆ The hypertextual contribution given by Pagerank eventually was ***not so intuitive*** as it seemed!
- ◆ And so even its analysis for spamdex is far from obvious!



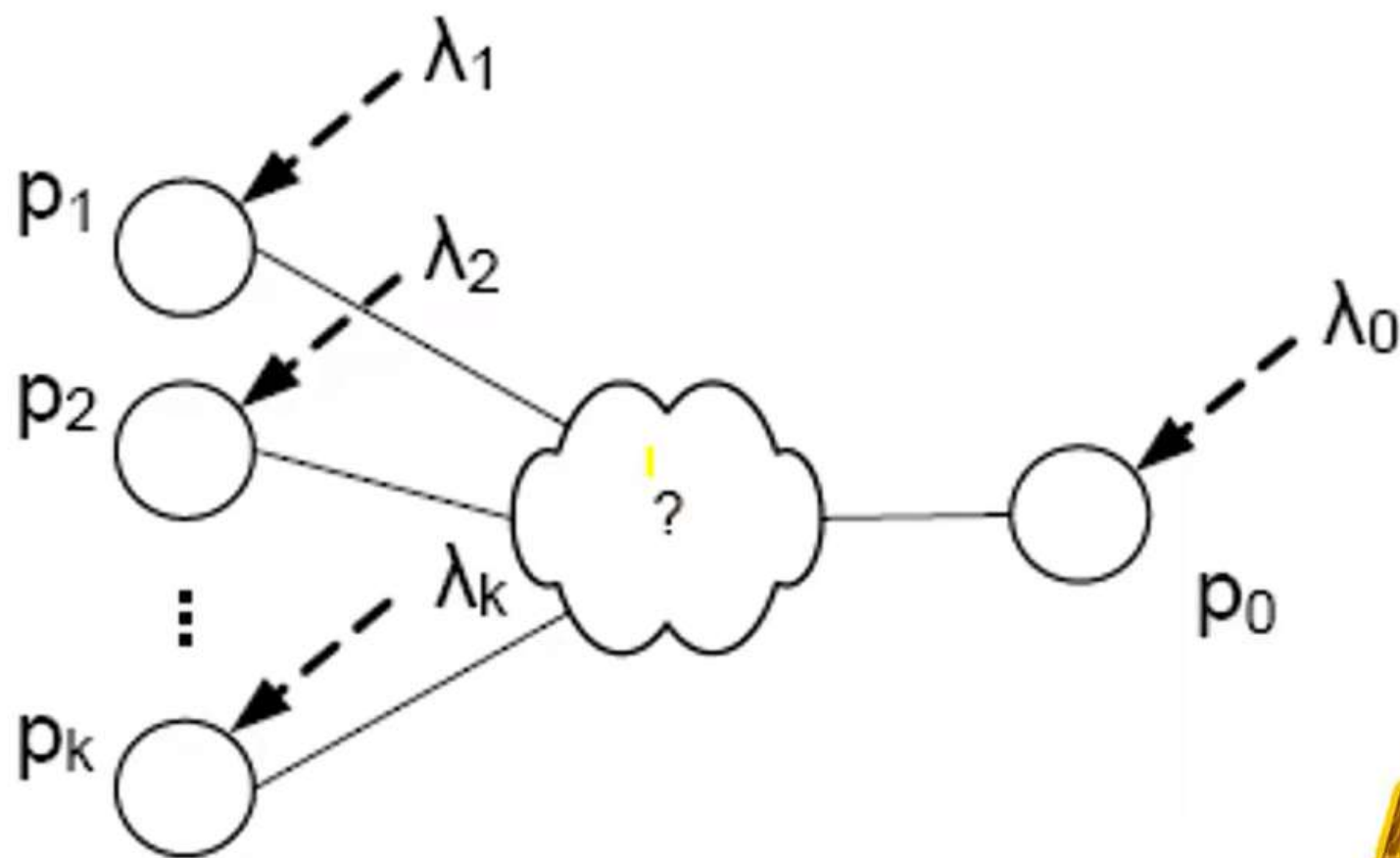
The Spam Farm

- ◆ A special link structure devoted to increase the hypertextual score

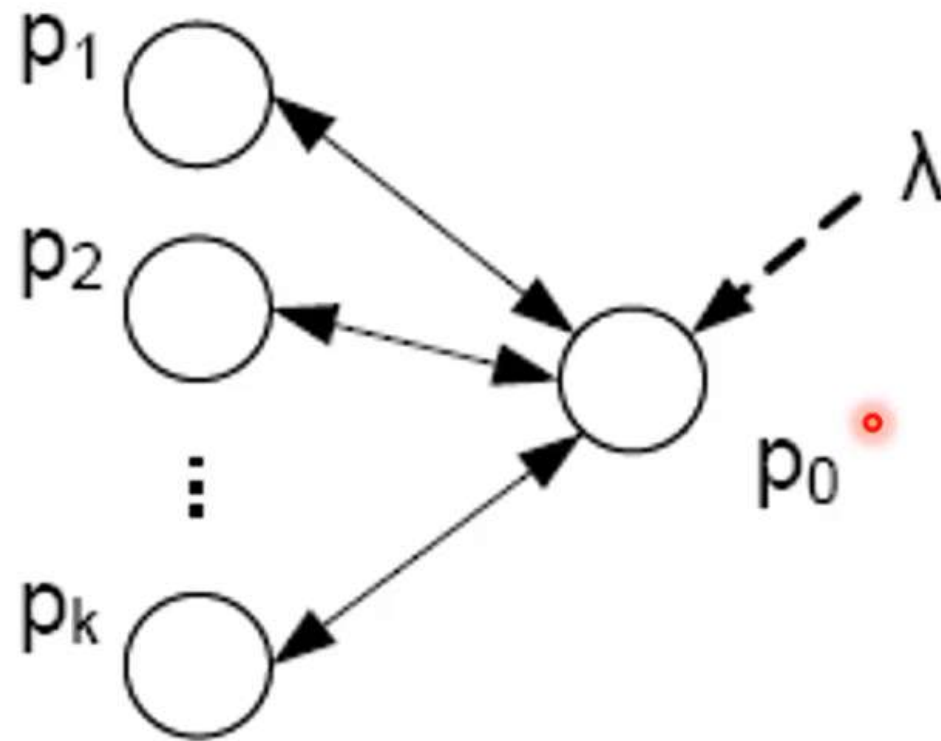




Structure of a generic farm

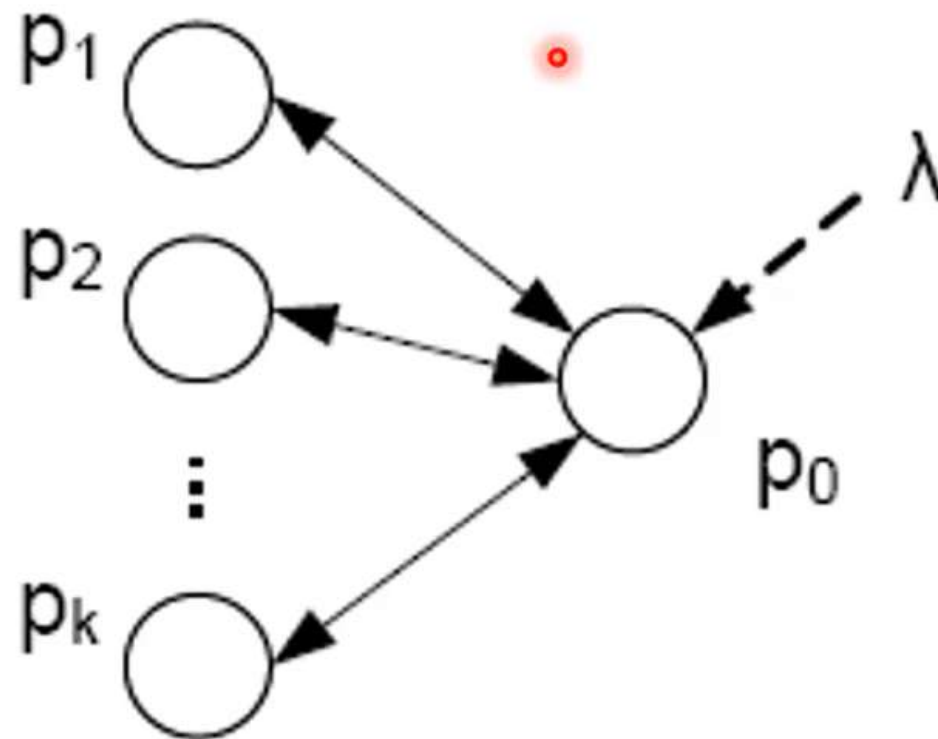


Optimal spam farm with a target page?



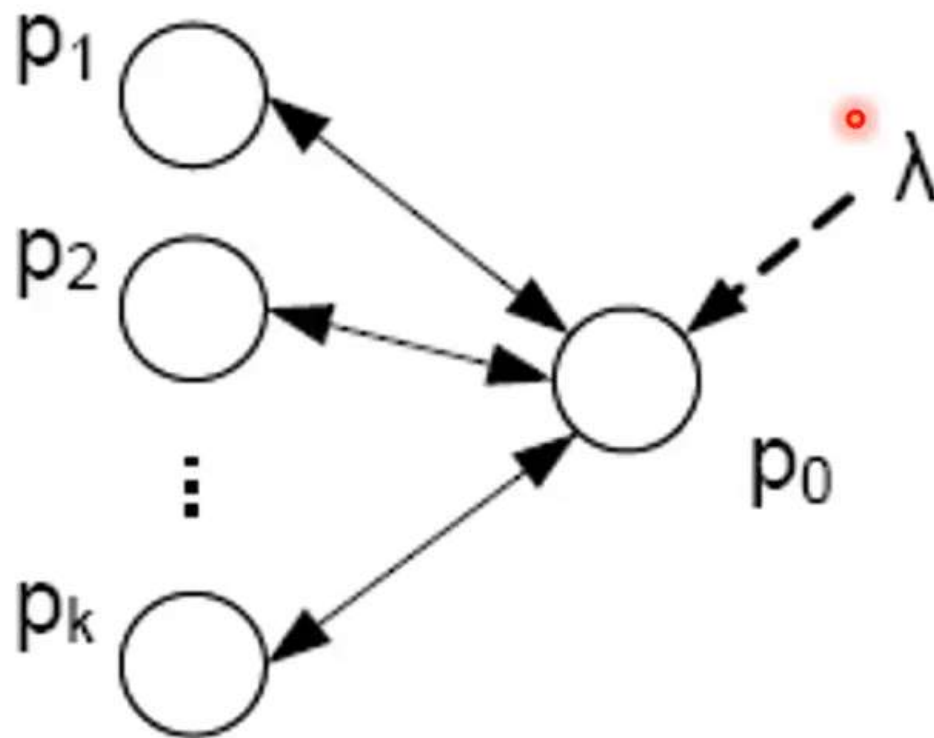
Good properties of this structure

- ◆ It uses the *least possible number of links*, while still keeping...



... a very important property

◆ ***Reachability***

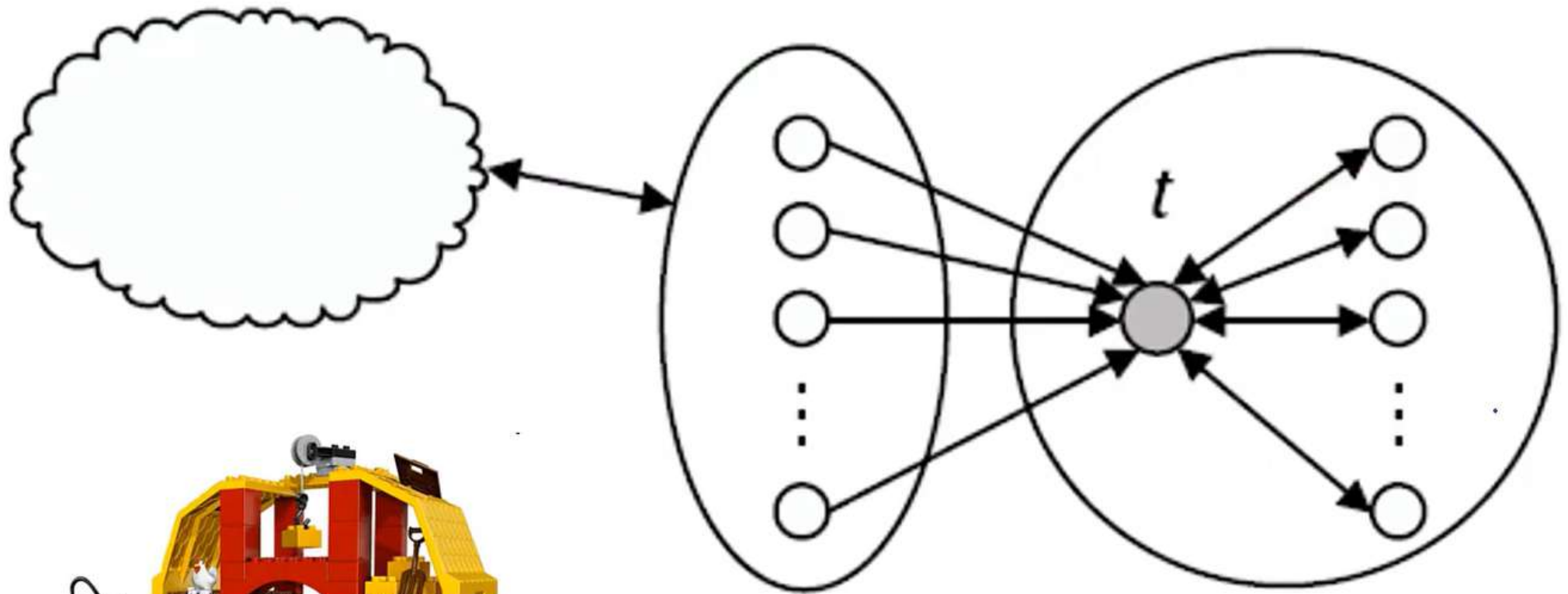


Optimal structure generalized

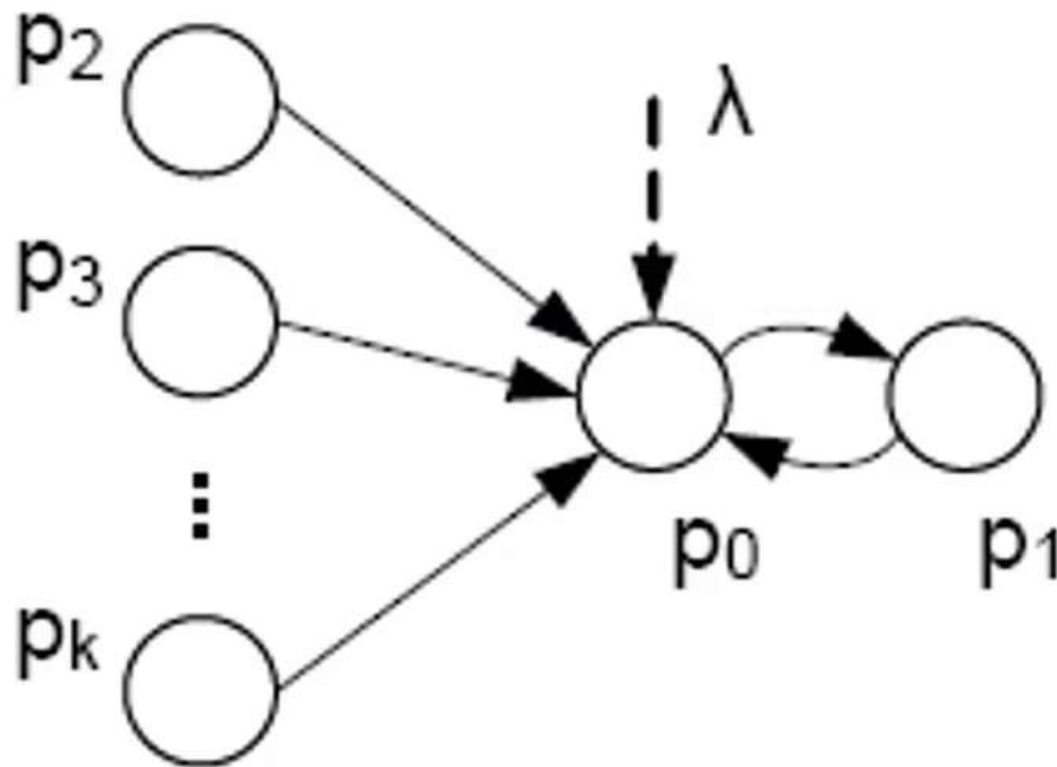
inaccessible

accessible

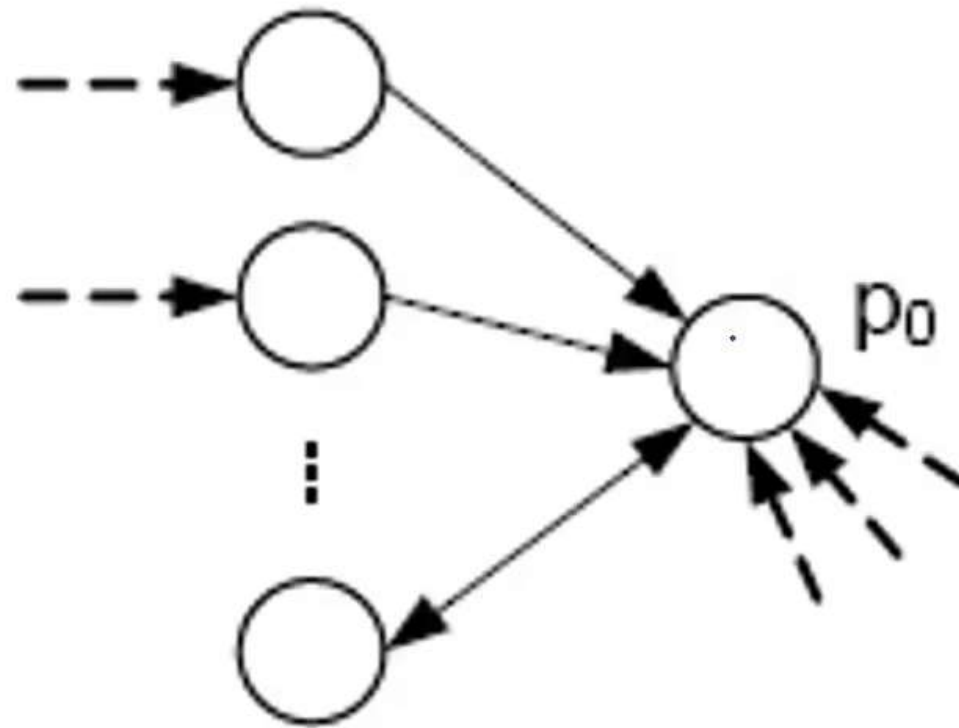
own



Another optimal spam farm (sacrificing reachability)



No reachability \rightarrow
will need some «patching»



Important aspect

- ◆ Are we alone? Cannot we join forces instead with an ***alliance***?
- ◆ The idea: join our spam farm with someone else's

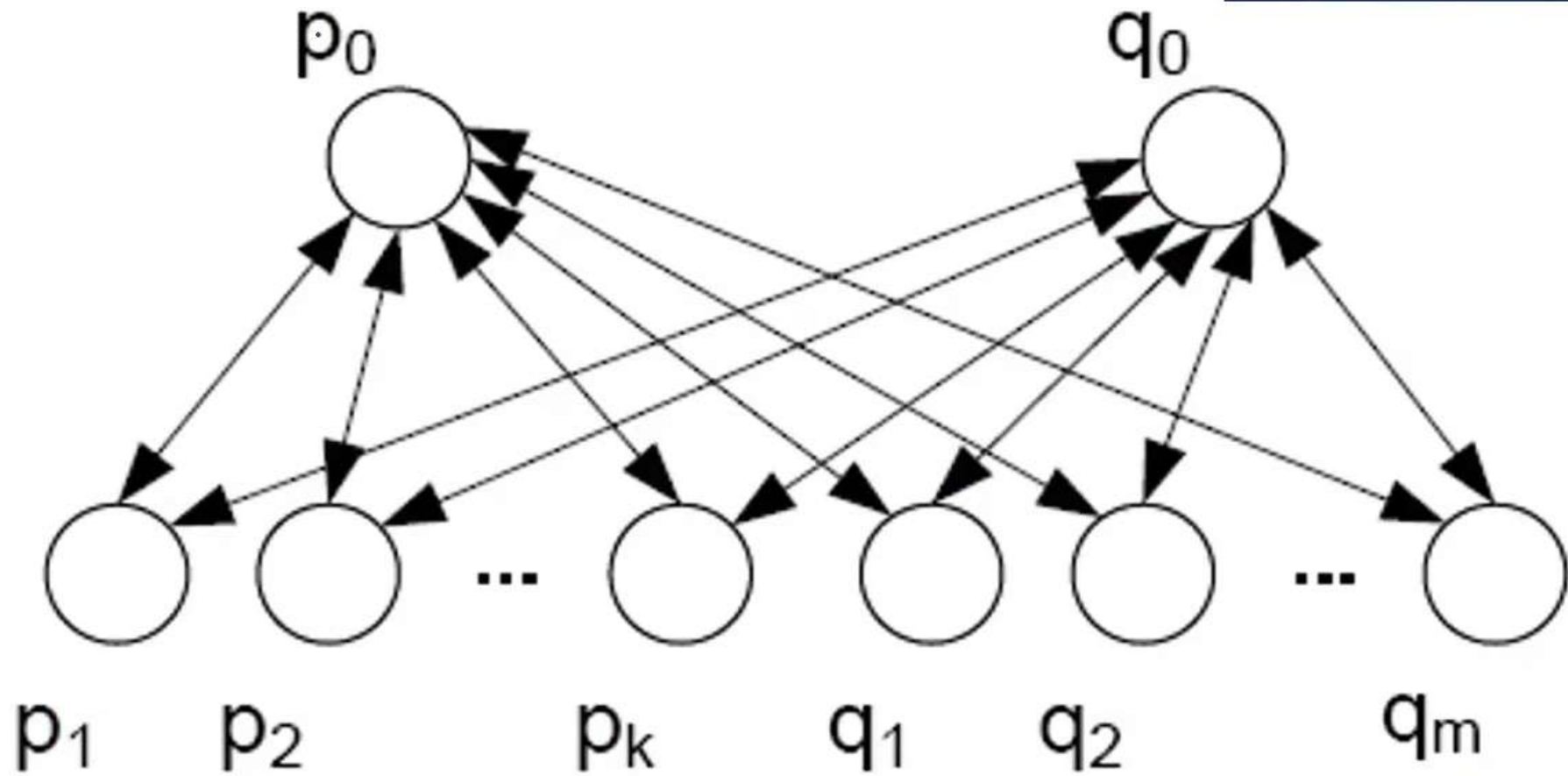


The Alliance Problem

◆ What are the best ways to do it?



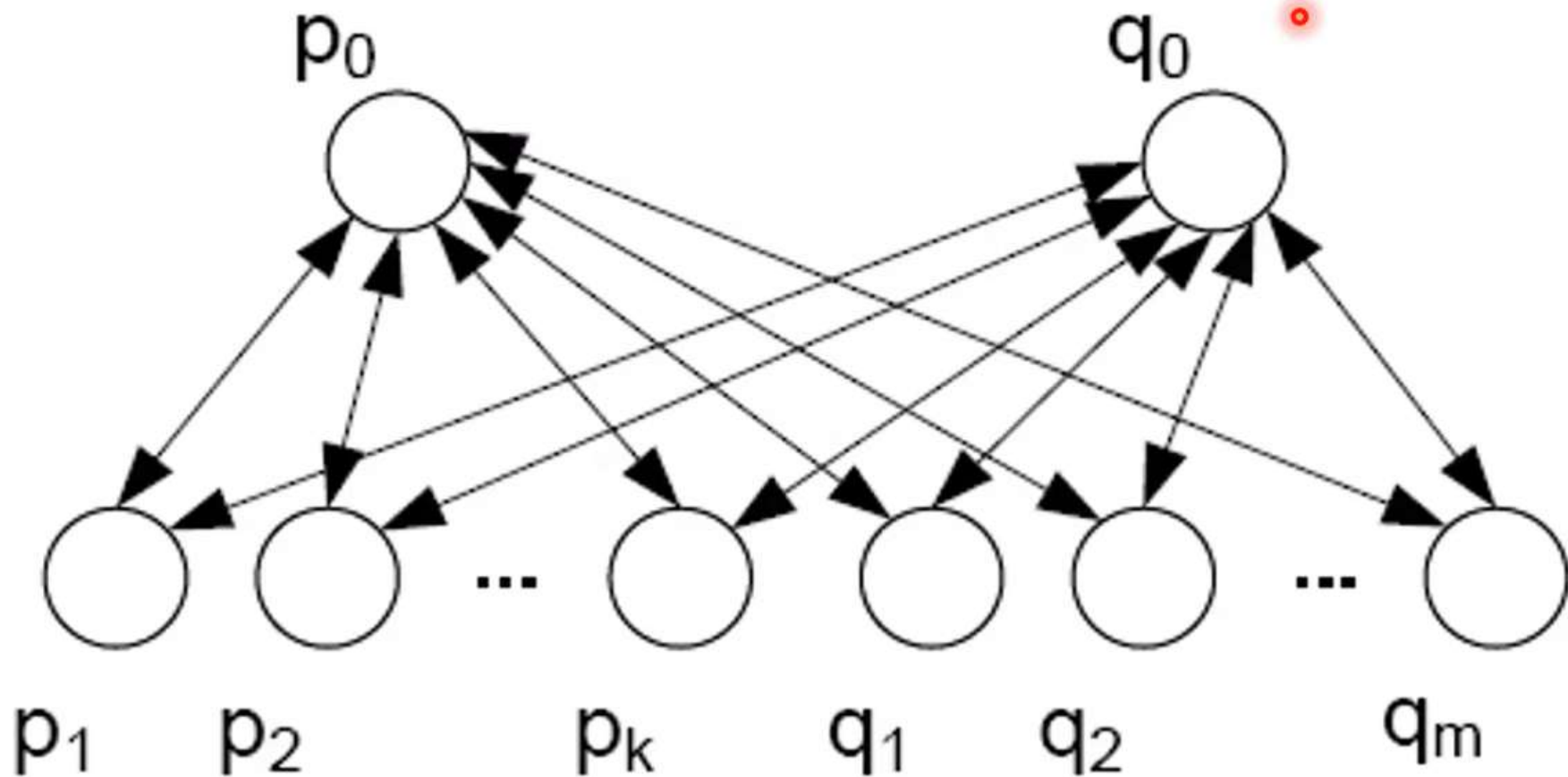
Scheme: «deep» alliance



How much pagerank?



◆ → the average of the two pageranks!

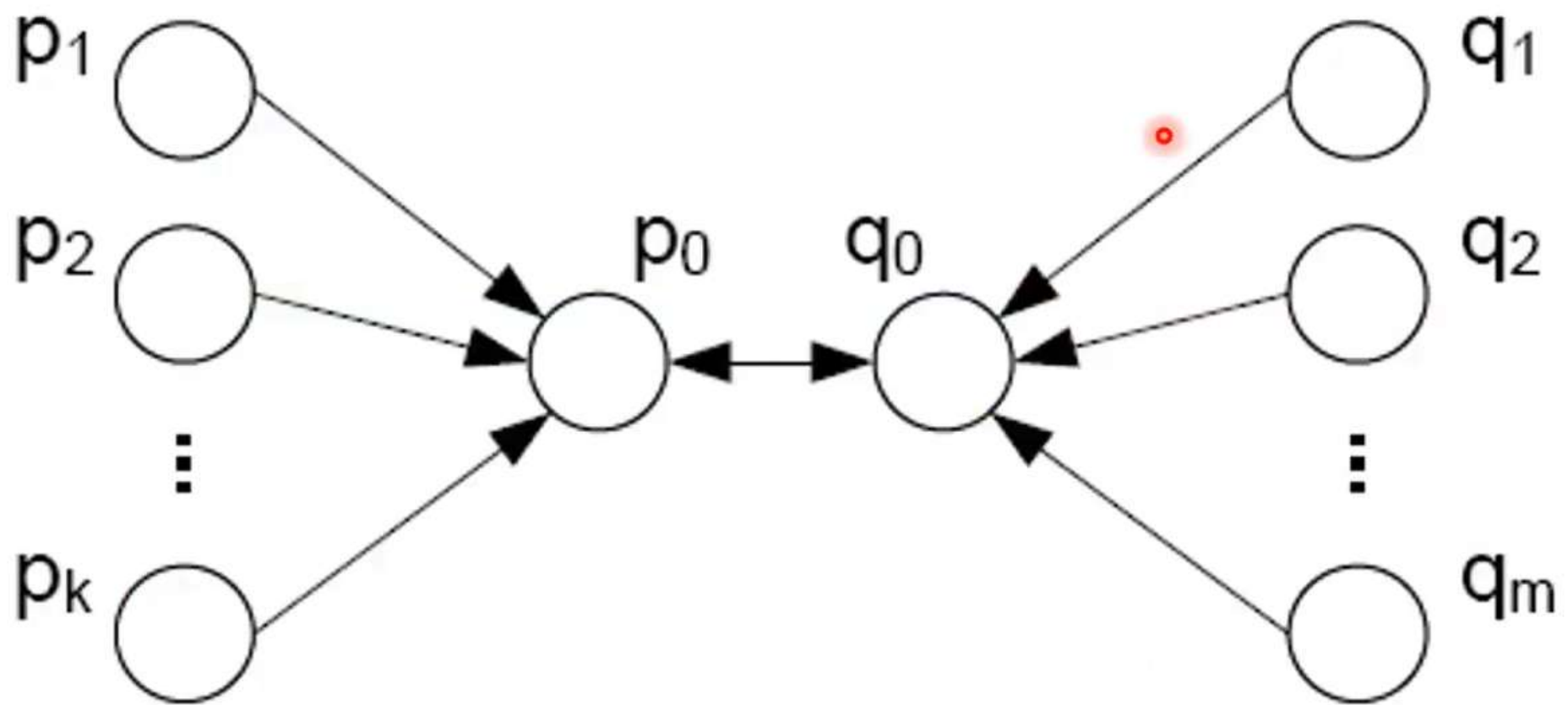


So...

- ◆ Useful alliance scheme to share the load and get a more ***robust*** configuration: we do our best and then we share goods and bads



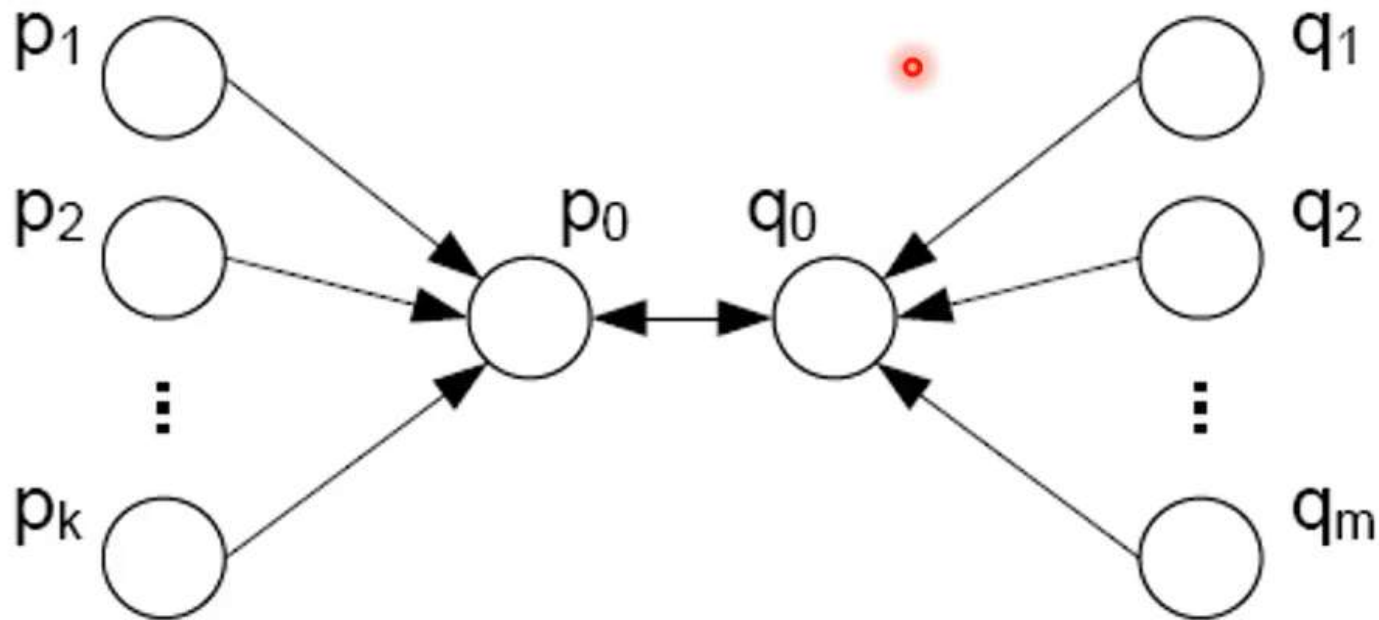
Scheme: «superficial» alliance



How much pagerank?

◆ **MORE THAN THE MAX BETWEEN THE TWO (!!!!)**

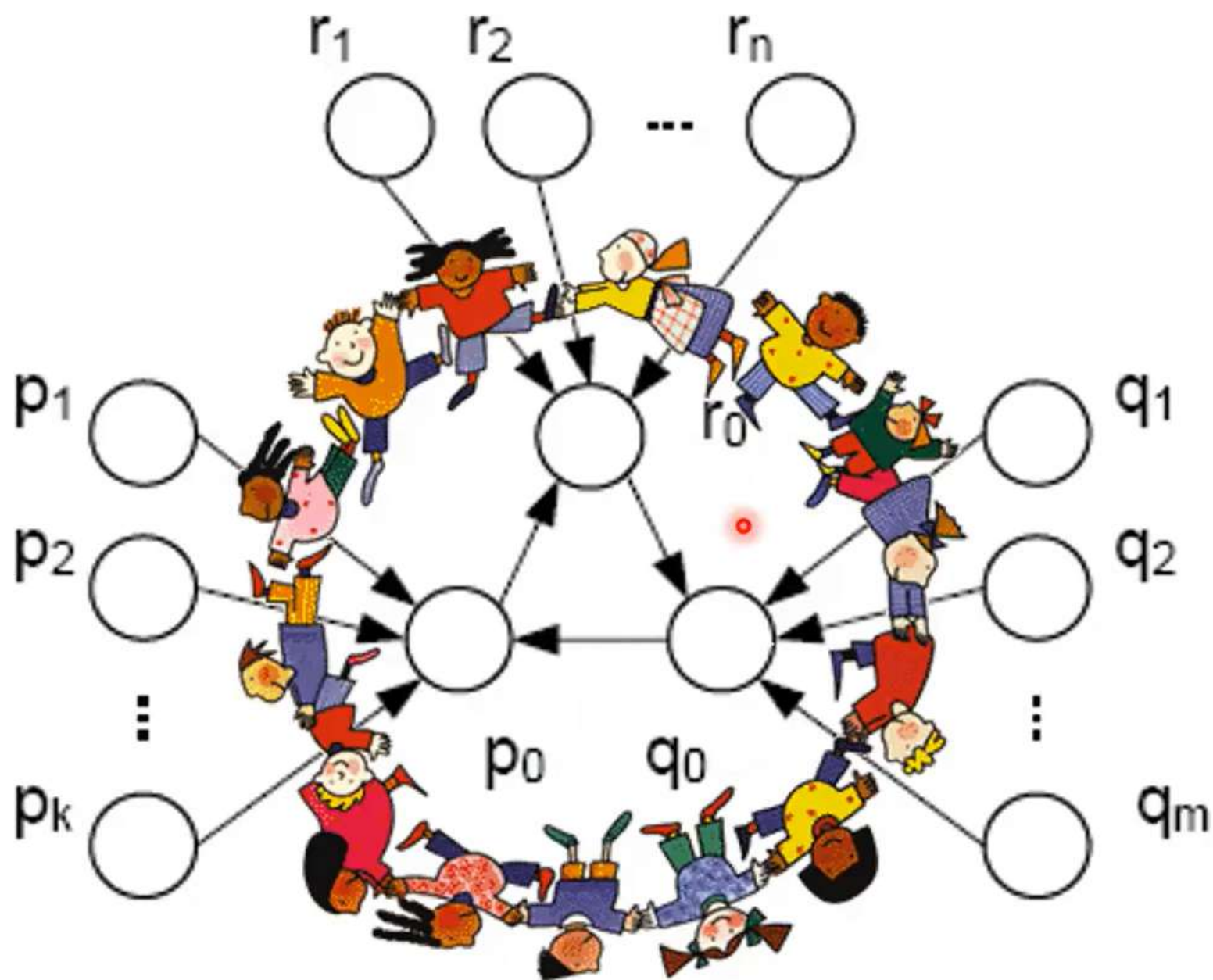
(bonus proportional to k and m)



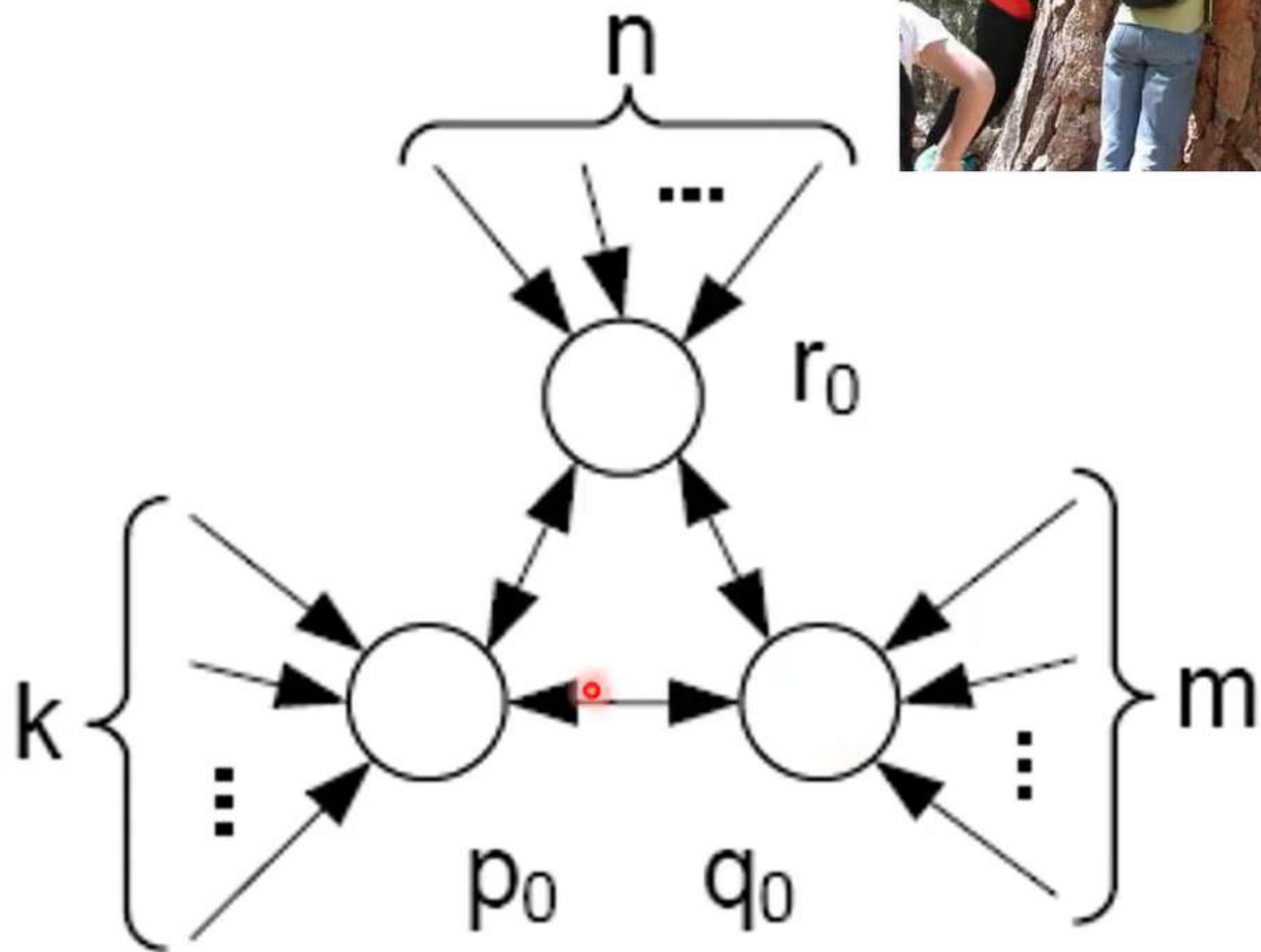
What about bigger alliances ?



Ring



Complete core



Observation...

◆ What about the poor search engines?





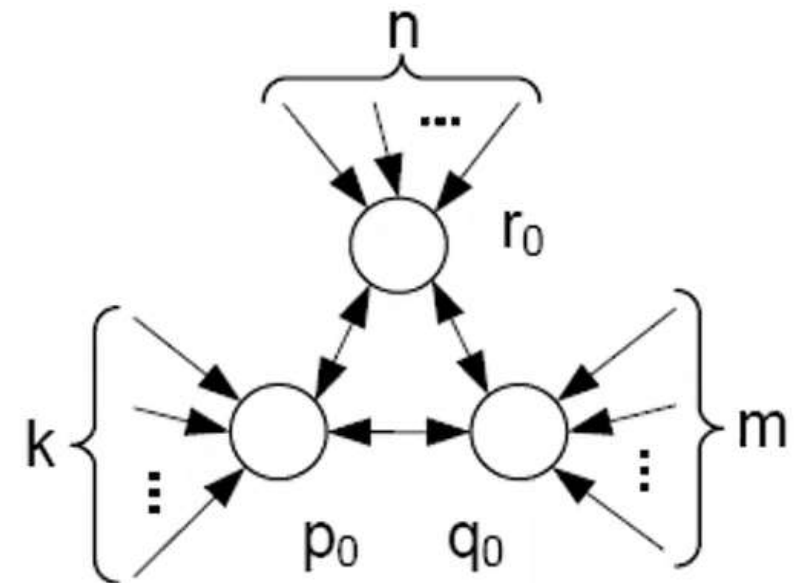
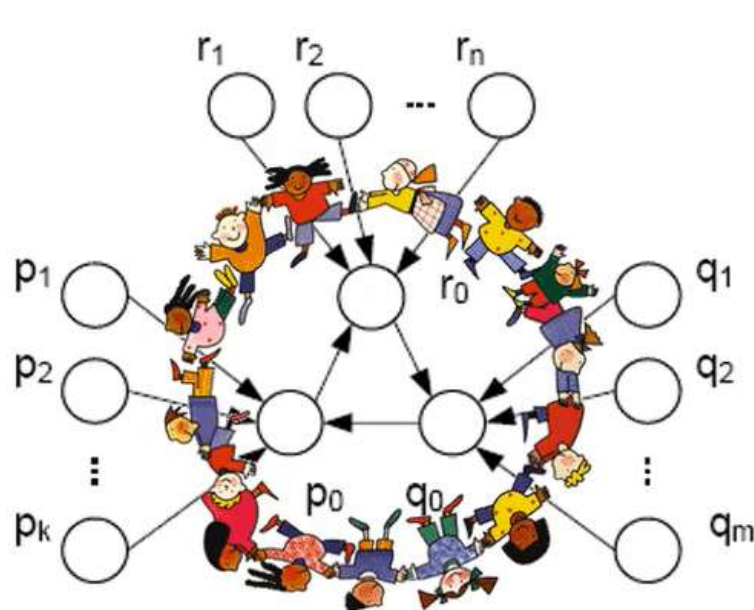
Countermeasures!

- ◆ In this case, it doesn't look so difficult: just find the ring / complete core structures among different sites 😊



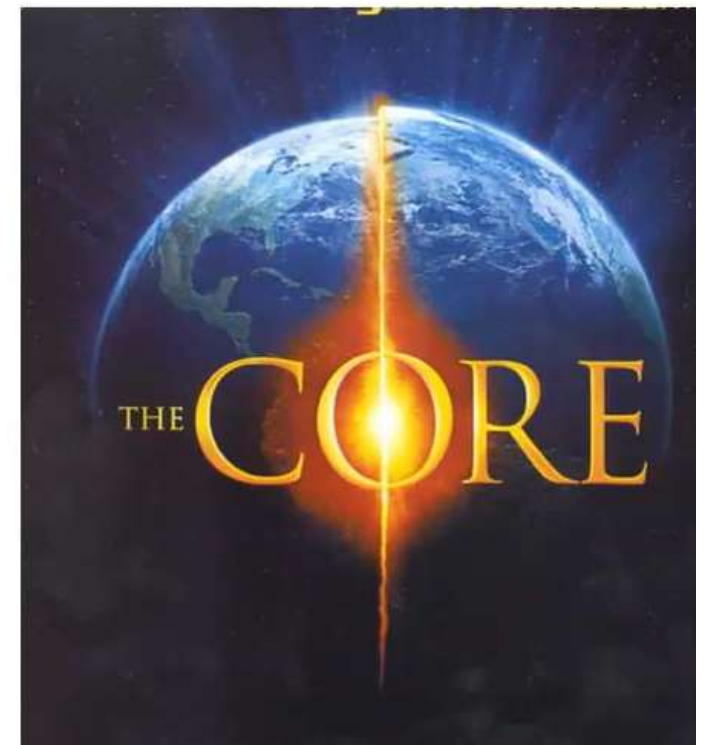
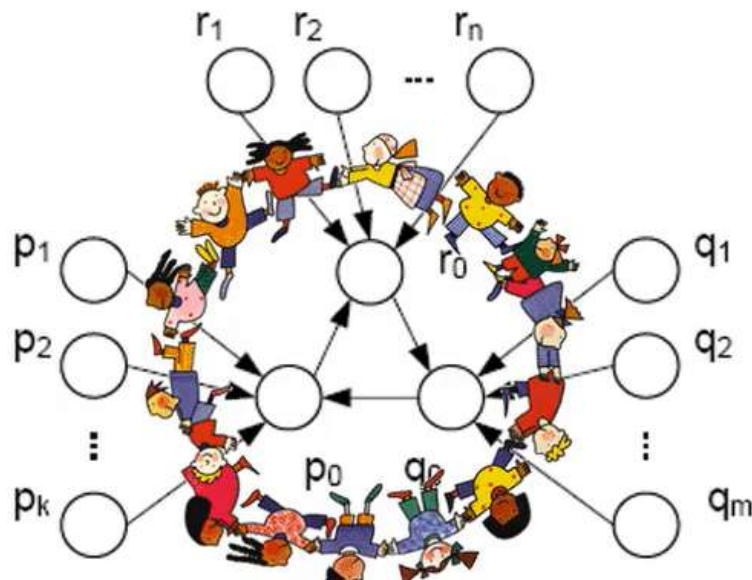
In reality...

- ◆ “Ring” and “complete core” are not the only two optimal structures
- ◆ There are other cores that give the same result



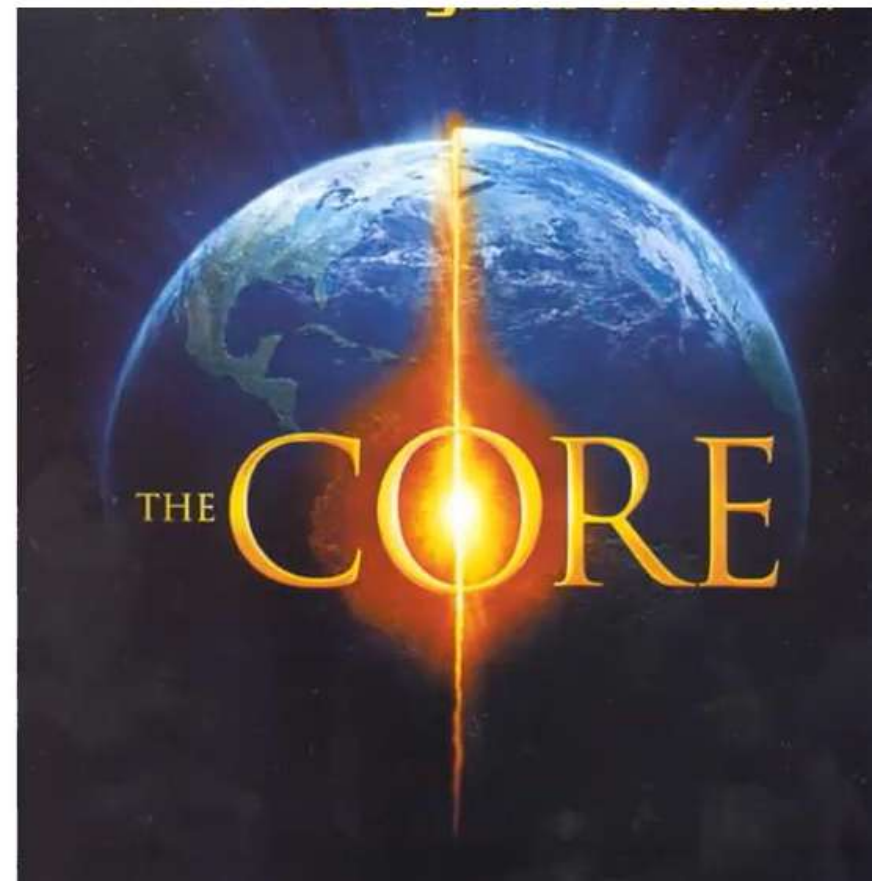
Other cores?

- ◆ We just need to have a ***strongly connected*** graph among the target pages (that is to say, from every page I can arrive to any other)



And so...

- ◆ ... how many strongly connected graphs there can be?
- ◆ If they are only a few ones, then the countermeasures will work well...
- ◆ So how many for an alliance (size N)?



The "A003030 sequence"...

- ◆ N=3 : 18
- ◆ N=4 : 1606
- ◆ N=5 : 565080
- ◆ N=6 : 734774776
- ◆ N=7 : 3523091615568
- N=8 : 63519209389664176
- N=9 : 4400410978376102609280
- N=10:1190433705317814685295399296
- N=11:1270463864957828799318424676767488
- ...

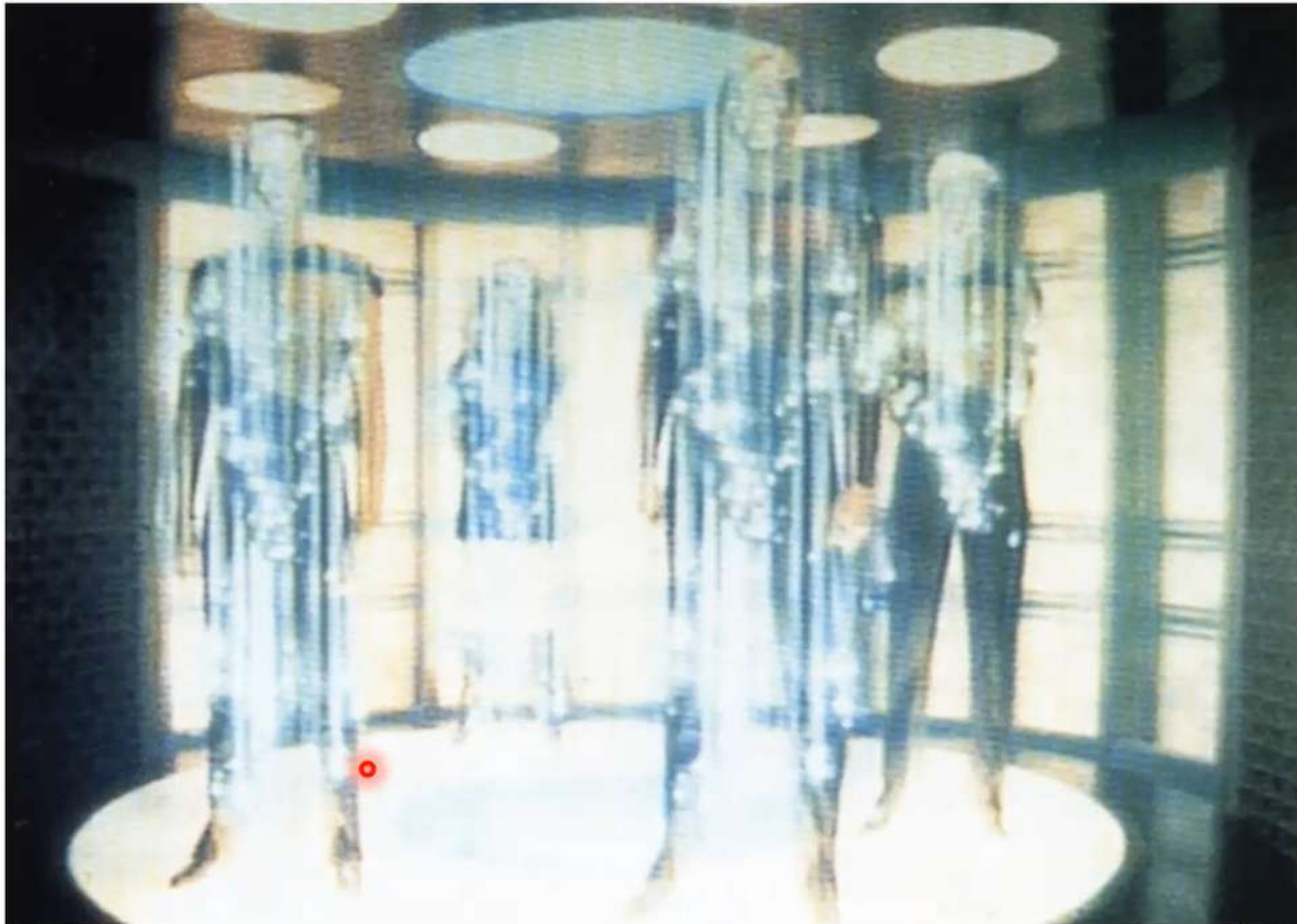


And so...

- ◆ ... Search engines are going to lose?
- ◆ They would need some special move...

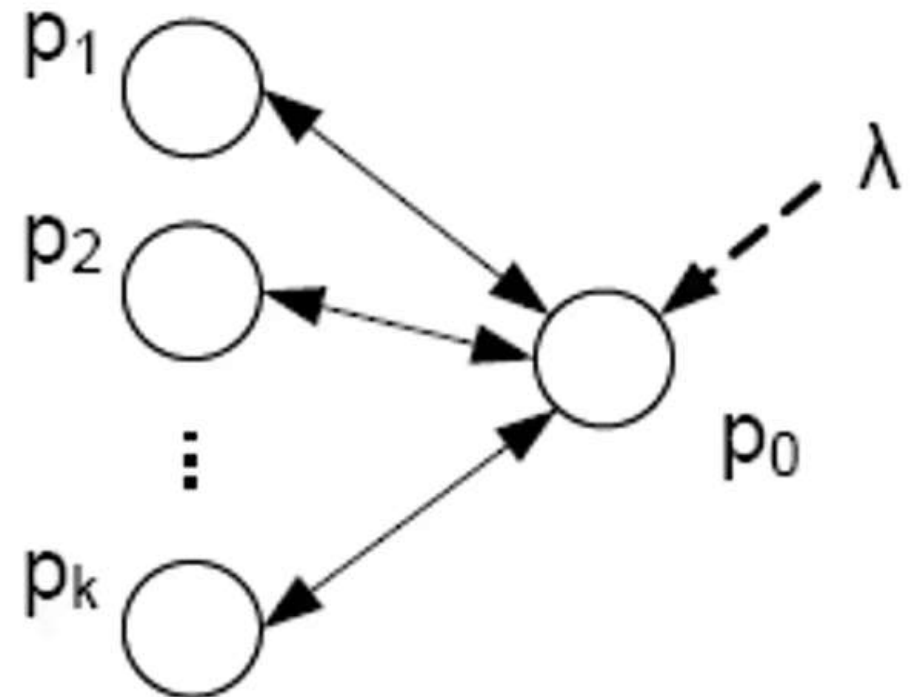
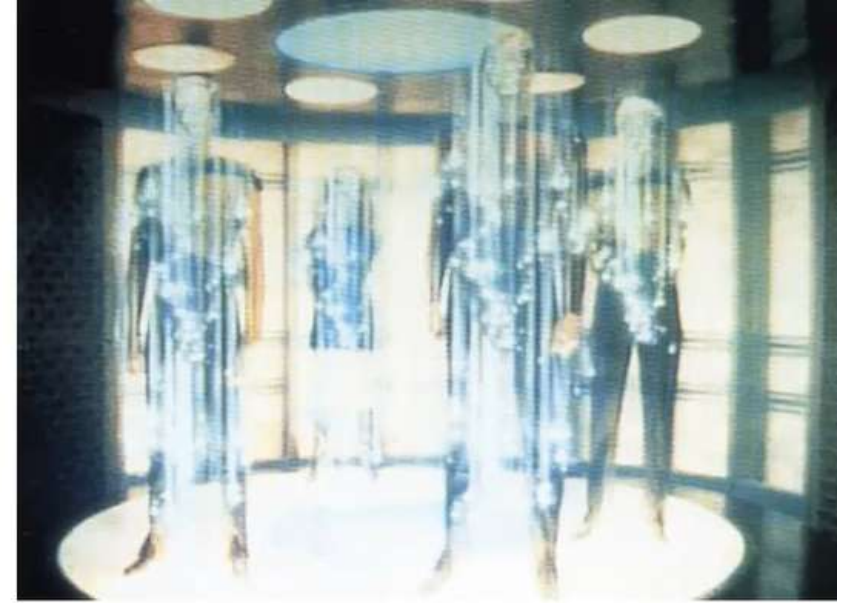


Let's see two of them...



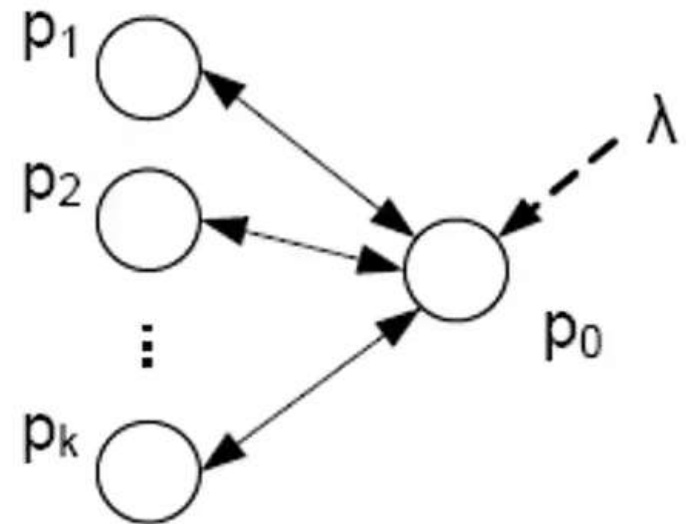
Example.....

- ◆ Remember the «teleport» component in pagerank
- ◆ What if we get rid of it?



This way...

- ◆ We could measure for a web page / web site what is the effect of teleportation
- ◆ A value called ***relative spam mass***
- ◆ → if it is too big, maybe something is wrong (too many «secondary» pages are offering contributions)



Result?

◆ Success rate to find spam:

◆ **95%-100%** 😊

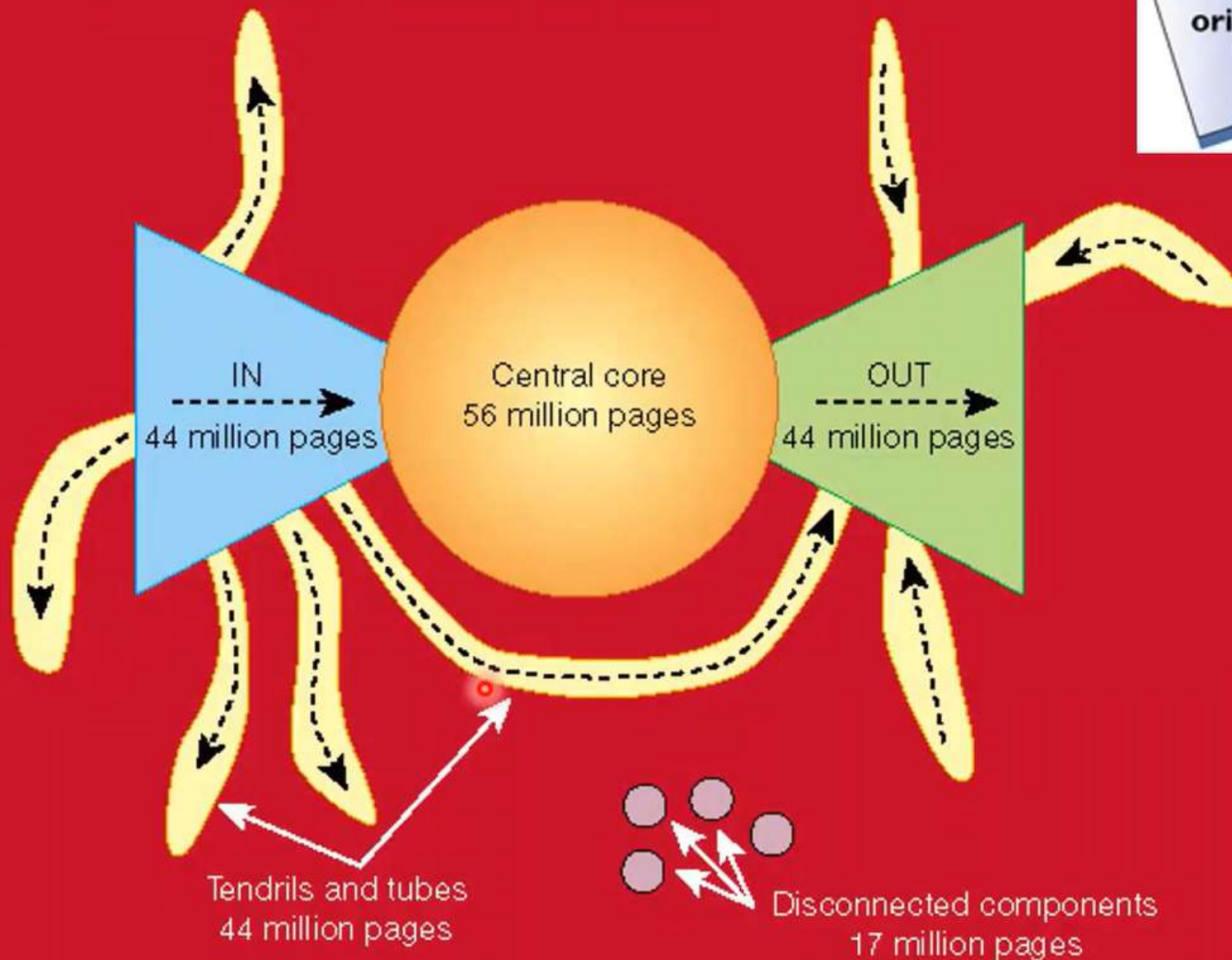
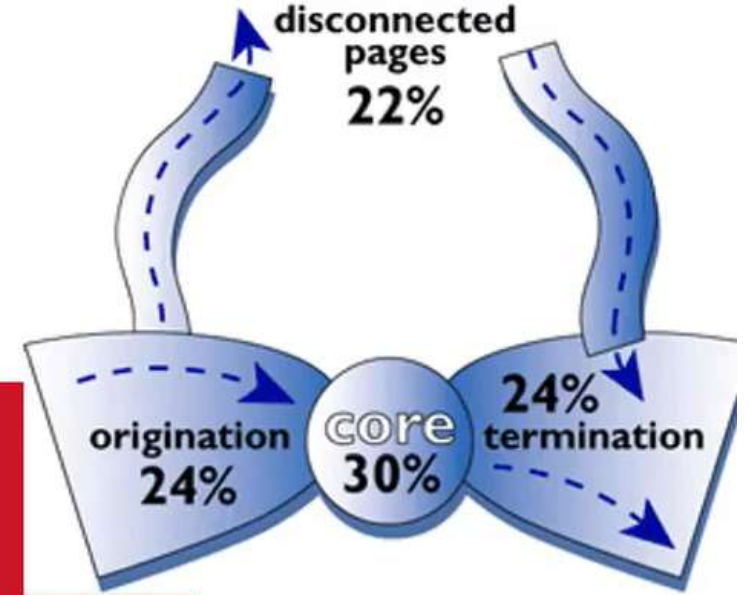
◆ (note: sufficient
but not necessary!)



Other example of countermeasure

- ◆ Does the web have a structure? Or is it just chaotic?





Super-Powerful countermeasure!

- ◆ Base idea: analyze the «shape» of a web site: if this is too much different from the average, then something is wrong
- ◆ All this (challenge) in an ***efficient way*** (!!)

