



Overview of Search Engine Spamming

Hector Garcia-Molina

Zoltán Gyöngyi

Roadmap

- What is search engine spam?
- What techniques do spammers use?
- Work at Stanford
- Challenges ahead

Example

The image shows a screenshot of a Google search results page for the query "kaiser pharmacy online". A yellow callout box with a black border points to the search term in the search bar. The search results are categorized under "Web" and show "Results 1 - 100 of about 175,000 for kaiser pharmacy online. (1.20 seconds)".

Search Bar: Web Images Groups News Froogle Local Scholar more »
kaiser pharmacy online Search Advanced Search Preferences

Web Results:

- Kaiser Health Insurance** (Sponsored Links)
www.kaisergrouphealth.com California only. Kaiser Permanente individual, family, and group plans
- Keiser College Online** (Sponsored Links)
ClassesUSA.com Attend College One Class At A Time At Keiser's eCampus. Learn More.
- Kaiser online pharmacy**
Best online pharmacy. Information on Kaiser online pharmacy and pet pharmacy online.
www.tak-news.com/online-pharmacy/kaiser-online-pharmacy.html - 40k -
Cached - Similar pages
- Online pharmacy schools**
kaiser online pharmacy and online pharmacy canada ... onlin phsmarcy rviews by kaiser
online pharmacy, kaiser online farmacie and canada pharmacy online. ...
www.tak-news.com/online-pharmacy/online-pharmacy-schools.html - 40k -
Cached - Similar pages
[More results from www.tak-news.com]
- Kaiser Pharmacy >> Online Pharmacy >> BugsBWear.com**
Home | Online Pharmacy | Page 7 | Kaiser Pharmacy. Pet Supplies · Beauty Products ·
Dental · Staying Healthy · Online Pharmacy. Kaiser Pharmacy ...
www.bugsbwear.com/pharmacy/term845.html - 9k - Cached - Similar pages
- Online Pharmacy >> Page 7 >> BugsBWear.com**
International Online Prescription Pharmacies · International Pharmacies Online ... Kaiser
Pharmacy · Kaiser S Drugstore · Kaisers Drugstore ...
www.bugsbwear.com/pharmacy/Page-7.html - 24k - Cached - Similar pages
- Combating Web Spam with TrustRank** (PDF)
File Format: PDF/Adobe Acrobat - View as HTML
kaiser pharmacy online. kaiser pharmacy online. Stop surfing! Only here. you can get
info about. kaiser pharmacy online. Click here to read info ...
www-db.stanford.edu/~zoltan/presentations/trustrank-vldb-2004-09-02.pdf - Similar pages
- Kaiser Pharmacy - Uncommonly Classical.com**
We're offer our online customers secure shopping with prices nobody beats! ... Your search
for "Kaiser Pharmacy" did not match any documents ...
www.uncommonlyclassical.com/Category/Pharmacy/Kaiser%20Pharmacy.html - 7k -
Cached - Similar pages
- Nground >> pharmacy**
Accutane Accutane - Actonel Actonel - Adderall online pharmacy Adderall online

Right Side (Sponsored Links):

- Prescriptions Online**
Are Online Prescriptions Safe?
Learn the risks of imported drugs.
www.buysafedrugs.info
- 50% off Online Pharmacies**
50% off Online Pharmacies
50% off drugs from Canada!
MedCenterCanada.com

Example

mp3

The screenshot shows a Google search for "online pharmacy". The search bar contains "online pharmacy" and the search button is labeled "Search". Above the search bar, navigation links include "Web", "Images", "Groups", "News", "Froogle", "mp3", and "more »". A yellow callout box labeled "mp3" points to the "mp3" link. Below the search bar, there is an advertisement for ZENEGRA (Sildenafil Citrate 100mg) with the text "TREAT ERECTILE DYSFUNCTION CHEAP! ONLY \$2.99! CHEAPEST PILLS ONLINE ORDER NOW!".

The search results are displayed under the heading "Web". The first result is "Carisoprodol Soma Watson 350mg/120tabs \$85" with a description: "Take Advantage of Our Prices without changing Quality! We Continue to provide the Best meds for Your Money! Cod, and Credit Cards Accepted Saturday delivery. Always Lower Prices, Always the Same Quality. U.S. Pharmacy, U.S. Meds. USA Orders Only!" and a URL: <http://www.mylifestylemeds.com/>. A yellow callout box labeled "Save today on Viagra, Lipitor, Zoloft, ..." points to this result.

The second result is "Internet's Pharmacy drug store" with a description: "Save today on Viagra, Lipitor, Zoloft and many other pharmaceutical drugs. FDA approved. Fedex overnight on all orders. Live online Help 24/7" and a URL: <http://www.bestpharmacydrugs.com>.

The third result is "Bay View Pharmacy: Your Reliable Online Pharmacy" with a description: "Buy FDA approved medications for pain relief, muscle relaxers and many more - FREE medical consultation. No Prior Prescription required Free Shipping." and a URL: <http://www.bayviewpharmacy.com>.

The fourth result is "Phentermine - Blow Out Sale!" with a description: "Best Weight Loss Pill - FDA approved, U.S. Licensed Physicians/ Pharmacist. Free Shipping and Customer Service ready to Help with your order. Orders In By 3:00 EST Shipped Same Day" and a URL: <http://www.YourDiscountDrugStore.com>. A yellow callout box labeled "Phentermine 90 Pills/\$119" points to this result.

The fifth result is "Phentermine 90 Pills/\$119" with a description: "Weekly Special: 90 pills (37.5mg). C.O.D., Free Shipping, U.S. Pharmacy, Free Consultation. Phone orders welcome. Orders In By 3:00 EST Shipped Same Day, Saturday Delivery." and a URL: <http://www.yourprescriptionforless.com>.

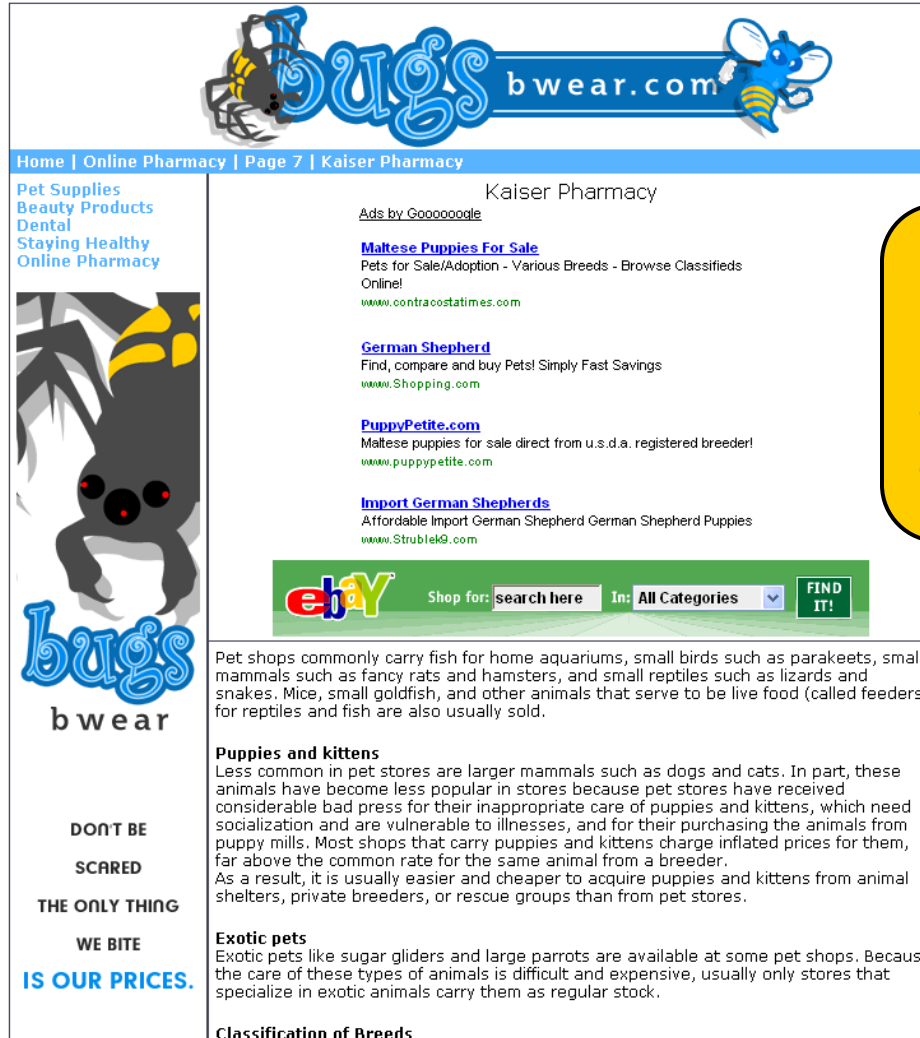
The sixth result is "Online Pharmacy Meds" with a description: "No prior prescription is needed. Fast, one page order form and FedEx overnight delivery of both brand name and generic drugs. Delivery to U.S. customers only." and a URL: <http://www.onlinepharmacymeds.com/?s=tw>.

The seventh result is "CostLessMedicine.com - Your Best Source For Meds" with a description: "Cost Less Medicine is your convenient, safe and private source for FDA approved prescription medications. We specialize in delivering quality medications at REALLY LOW PRICES! FREE Shipping by FedEx!" and a URL: <http://www.costlessmedicine.com>.

The eighth result is "Phentermine Sale! 37.5mg 90 pills \$119.00" with a description: "From A Company You Have Learned to Love and Trust! We are committed to providing..." and a URL: [http://www.yourprescriptionforless.com](#).

On the right side of the search results, there is a "Sponsored Links" section. The first link is "No Prior Script req'd + Free Shipping" with a description: "Low Price Prescriptions on soma, phentermine, fioricet, tramadol and much more! Overnight shipping, discreet packaging, and Free online doctors consultation." and a URL: www.superiorhealthpharmacy.com. The second link is "Best Online Pharmacy + FREE Shipping" with a description: "U.S FDA approved prescription medications. No hidden fees, no consultation fees. Free shipping. Great Prices. Order on line or call us to place your order." and a URL: www.stronghealthpharmacy.com. The third link is "Free Saturday Delivery - 5% Off Reorder" with a description: "FDA approved meds of the highest quality-US licensed Doctors and Pharmacies- Don't be fooled by those foreign imitation meds sold by our competitors-Free Sat Delivery-5% Discount On Reorders." and a URL: fastfedexmeds.com/index.php?c=7s&ck=index.

Example



The screenshot shows the homepage of bugs bwear.com. At the top is the logo with a bee and the text 'bugs bwear.com'. Below the logo is a navigation bar with links for 'Home', 'Online Pharmacy', 'Page 7', and 'Kaiser Pharmacy'. The main content area is divided into two columns. The left column contains a vertical list of links: 'Pet Supplies', 'Beauty Products', 'Dental', 'Staying Healthy', and 'Online Pharmacy'. Below these links is a large graphic of a bee with the text 'bugs bwear' underneath. The right column features a search bar with the text 'Shop for: search here In: All Categories' and a 'FIND IT!' button. Below the search bar are several pet-related advertisements, including 'Maltese Puppies For Sale', 'German Shepherd', 'PuppyPette.com', and 'Import German Shepherds'. At the bottom of the page, there is a section titled 'Classification of Breeds'.

Home | Online Pharmacy | Page 7 | Kaiser Pharmacy

bugs bwear

Pet Supplies
Beauty Products
Dental
Staying Healthy
Online Pharmacy

Kaiser Pharmacy
Ads by Google

Maltese Puppies For Sale
Pets for Sale/Adoption - Various Breeds - Browse Classifieds Online!
www.contracostatimes.com

German Shepherd
Find, compare and buy Pets! Simply Fast Savings
www.Shopping.com

PuppyPette.com
Maltese puppies for sale direct from u.s.d.a. registered breeder!
www.puppypette.com

Import German Shepherds
Affordable Import German Shepherd German Shepherd Puppies
www.Strubler9.com

Shop for: In:

Pet shops commonly carry fish for home aquariums, small birds such as parakeets, small mammals such as fancy rats and hamsters, and small reptiles such as lizards and snakes. Mice, small goldfish, and other animals that serve to be live food (called feeders) for reptiles and fish are also usually sold.

Puppies and kittens
Less common in pet stores are larger mammals such as dogs and cats. In part, these animals have become less popular in stores because pet stores have received considerable bad press for their inappropriate care of puppies and kittens, which need socialization and are vulnerable to illnesses, and for their purchasing the animals from puppy mills. Most shops that carry puppies and kittens charge inflated prices for them, far above the common rate for the same animal from a breeder. As a result, it is usually easier and cheaper to acquire puppies and kittens from animal shelters, private breeders, or rescue groups than from pet stores.

Exotic pets
Exotic pets like sugar gliders and large parrots are available at some pet shops. Because the care of these types of animals is difficult and expensive, usually only stores that specialize in exotic animals carry them as regular stock.

Classification of Breeds

**DON'T BE
SCARED
THE ONLY THING
WE BITE
IS OUR PRICES.**

Pet shops commonly carry fish for home aquariums, small birds such as parakeets, small mammals such as fancy rats and hamsters...

Example

UNCOMMONLY CLASSICAL.COM
ONLINE SHOPPING EXTRAVEGANZA!

NAVIGATE

- Lawyers
- Loans
- Magazines
- Mortgage
- Nurse
- Pharmacy
- Posters And Crafts
- Ringtones
- Shoes
- Stock Market
- Sunglasses
- Theater
- Viagra
- Video Games

HOT PRODUCTS

ZERO DOWNTIME
Host your website on 5 machines!
\$7.77 PER MONTH

ABOUT US

WHEN PEOPLE THINK ABOUT SHOPPING, THEY THINK OF GOING TO A STORE, BUT HOW ABOUT IF THE STORE CAME TO YOU! THAT'S RIGHT, THROUGH SOMETHING WE LIKE TO CALL THE "INTERNET", WE HAVE REACHED YOU IN YOUR HOME!

GO SHOPPING NOW!

TAKE A LOOK AT SOME OF OUR TOP SELLERS!

Ads by Goooooogle

Compounded Medications
Quality drugs, nationwide delivery Good prices, great customer service
McGuffCompoundingPharmacy.com

Pharmacist-CA
We recruit pharmacists for CA jobs. We represent the top CA employers.
www.pharmacist-ca.com

Pharmacy
Safe, Quality, Prescription Drugs Licensed Canadian Pharmacy
www.CanPharm.com

Online Pharmacy
Save on thousands of drugs. Local Pharmacy quality drugs
www.usamedsonline.com

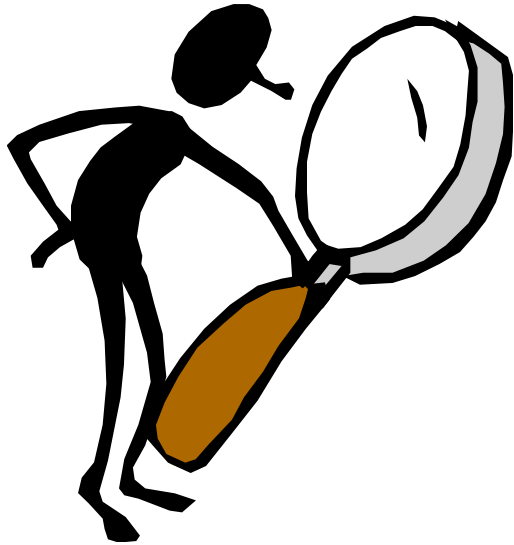
ARTICLES

If you're looking for PHARMACY, **Click Here!**
Pharmacy is the profession of compounding and dispensing medication. More recently, the term has come to include other services related to patient care including clinical practice, medication review, drug information, etc. Some of these new roles are now mandated by law in various legislatures. Pharmacists, therefore, are the primary health professionals who optimise

Lawyers
Loans
Mortgage
Ringtones
Viagra

Pharmacy is the profession of compounding and dispensing medication. More recently, the term has come to include other services...

Definition



So... who does what?

Definition



So... who does what?

Spamming

deliberate human action

Definition



So... who does what?

Spamming

deliberate human action

meant to trigger unjustifiably high
ranking

Definition



So... who does what?

Spamming

deliberate human action

meant to trigger unjustifiably high

ranking

importance
(*global*)

relevance
(*query-dependent*)

Monetization

Why?

- Better ranking = higher click-through rate
 - Search engine optimization
 - Affiliate spam
 - Advertisement spam

[Compare Canon Prices](#) - search.ActivShopper.com
Compare Prices For Canon Cameras In The USA Before You Buy!

[Ads by Goooooogle](#)

[Study online at MusicianUniversity.com](#)
[Study audio, music and instruments courses](#)
[Start your course with 25% off ONLY Today](#)

Canon Digital Rebel XT 8MP Digital SLR Camera with EF-S 18-55mm f3.5-5.6 Lens (Black)

Features

- Powered by rechargeable Lithium-ion battery (included, with charger)
- 8.0-megapixel CMOS sensor captures enough detail for photo-quality 16 x 22-inch prints
- Includes Canon's EF-S 18-55mm, f3.5-5.6 zoom lens
- DIGIC II Image Processor provides fast, accurate image processing; captures images at a rate of up to 3 frames per second
- Fast start-up time-- 2 seconds

Canon Digital Rebel XT 8MP Digital SLR Camera with EF-S 18-55mm f3.5-5.6 Lens (Black) Reviews

This is one awesome camera

🎵 🎵 🎵 🎵 🎵 🎵 🎵 🎵 (10 out of 10)

I bought from amazon after being let down by Dell. It arrived next day and on the weekend i gave it a good workout at a local airshow. I took just under 1000 photos and was amazed by how well the shots came out in sport mode. Very quick focus and spot on. excellent camera and well built , not small as some people have noted, works for me...



Ready To Buy Online Now?

List Price: ~~\$1,499.99~~
Online Price at this moment: **Too Low To Display \$1011.67**
Buy used: \$915.00
\$440.00

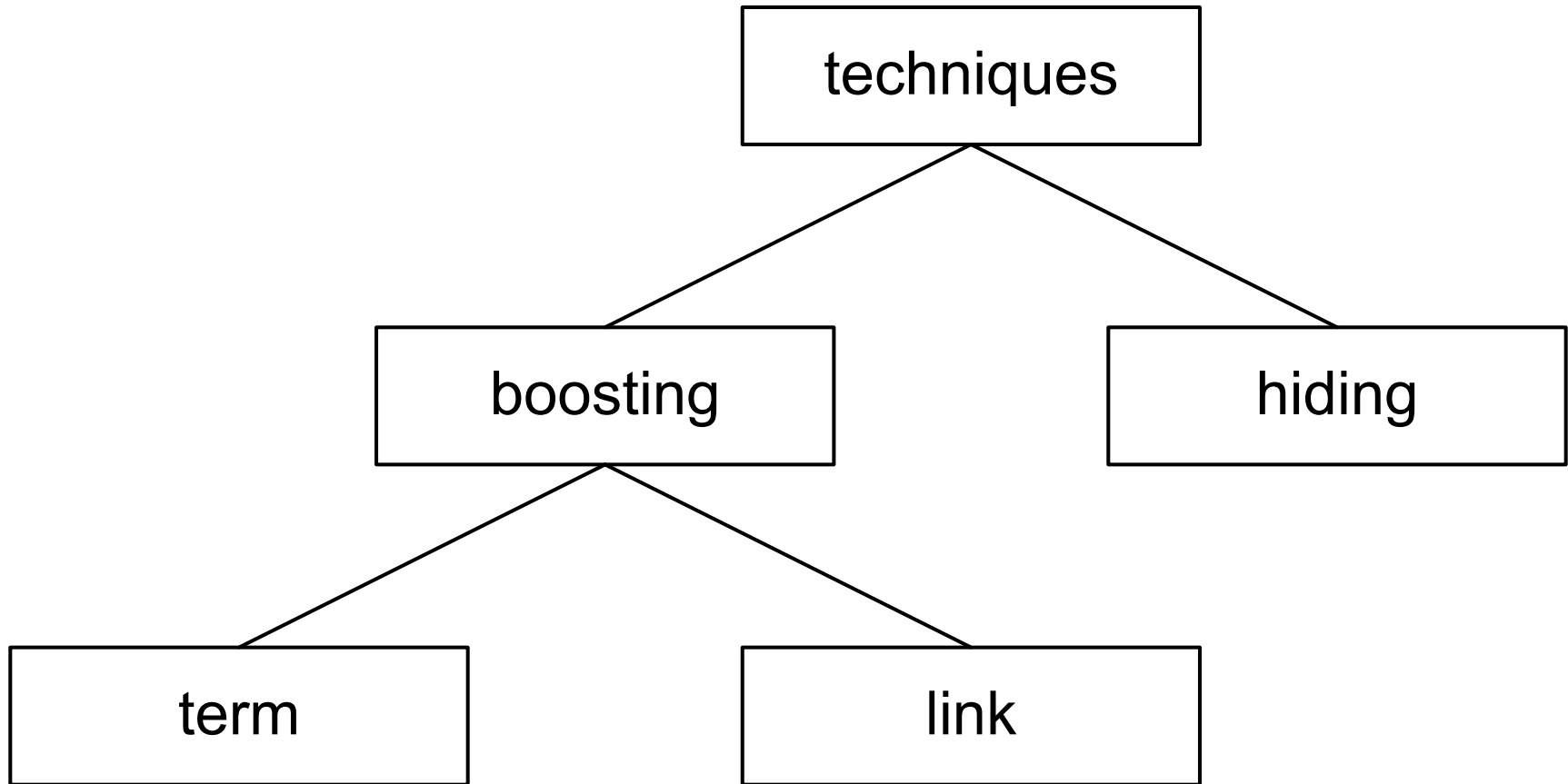
[Buy, See Prices](#)

Canada | UK | Germany/Europe

CRMAV.com Pro Audio
MusiciansNews.com
DJMusicEquipment.com
GuitarsMusicGear.com
GuitarSheetMusicTabs.com
Digital Pianos Keyboards
Live Sound PA Systems
RecordingMixingMastering.com
DigitalCamerasPhotography.com
DigitalVideoCameraEditing.com
ShoesClothesFashion.com
CarsMotorsTrucks.com
StayLimy.com Health/Sports
PC4D.com
Computers/Technology
House Home Furniture

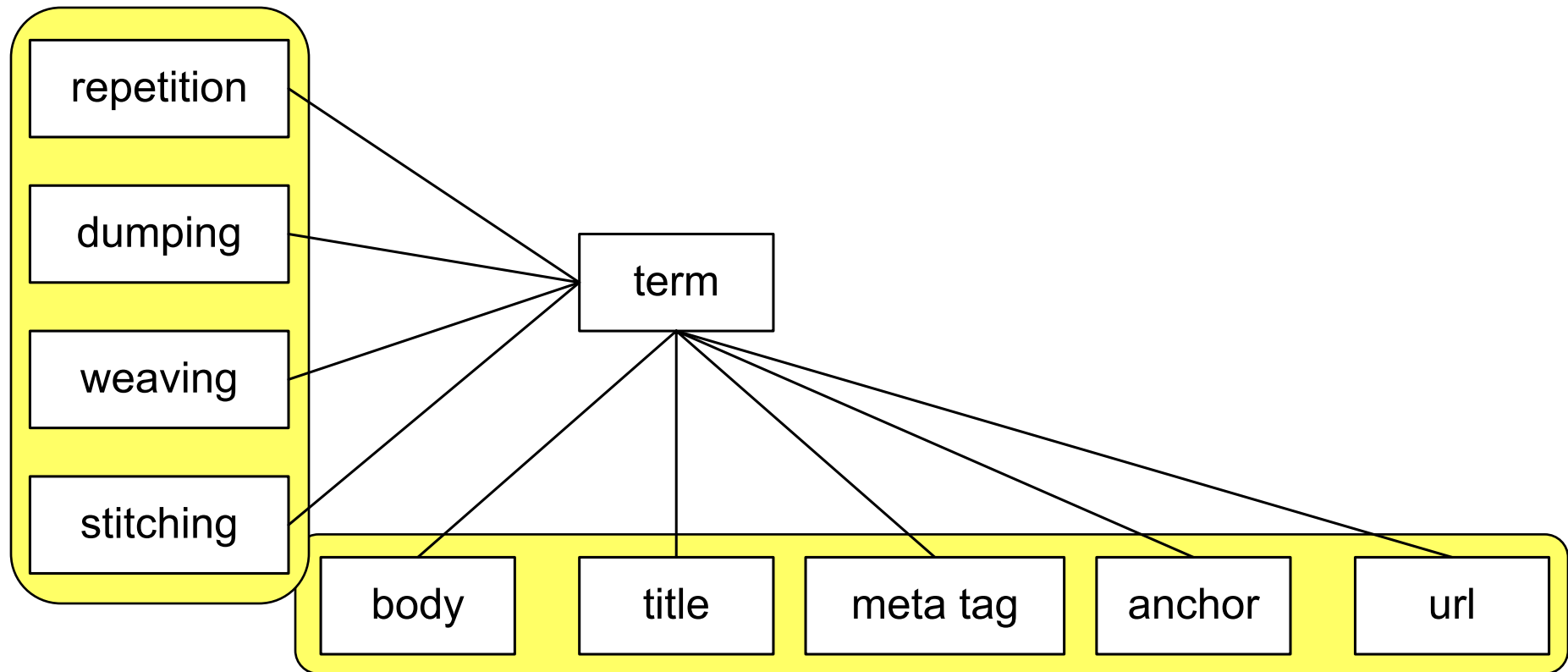
Techniques

How?



Techniques / Boosting / Term

how?

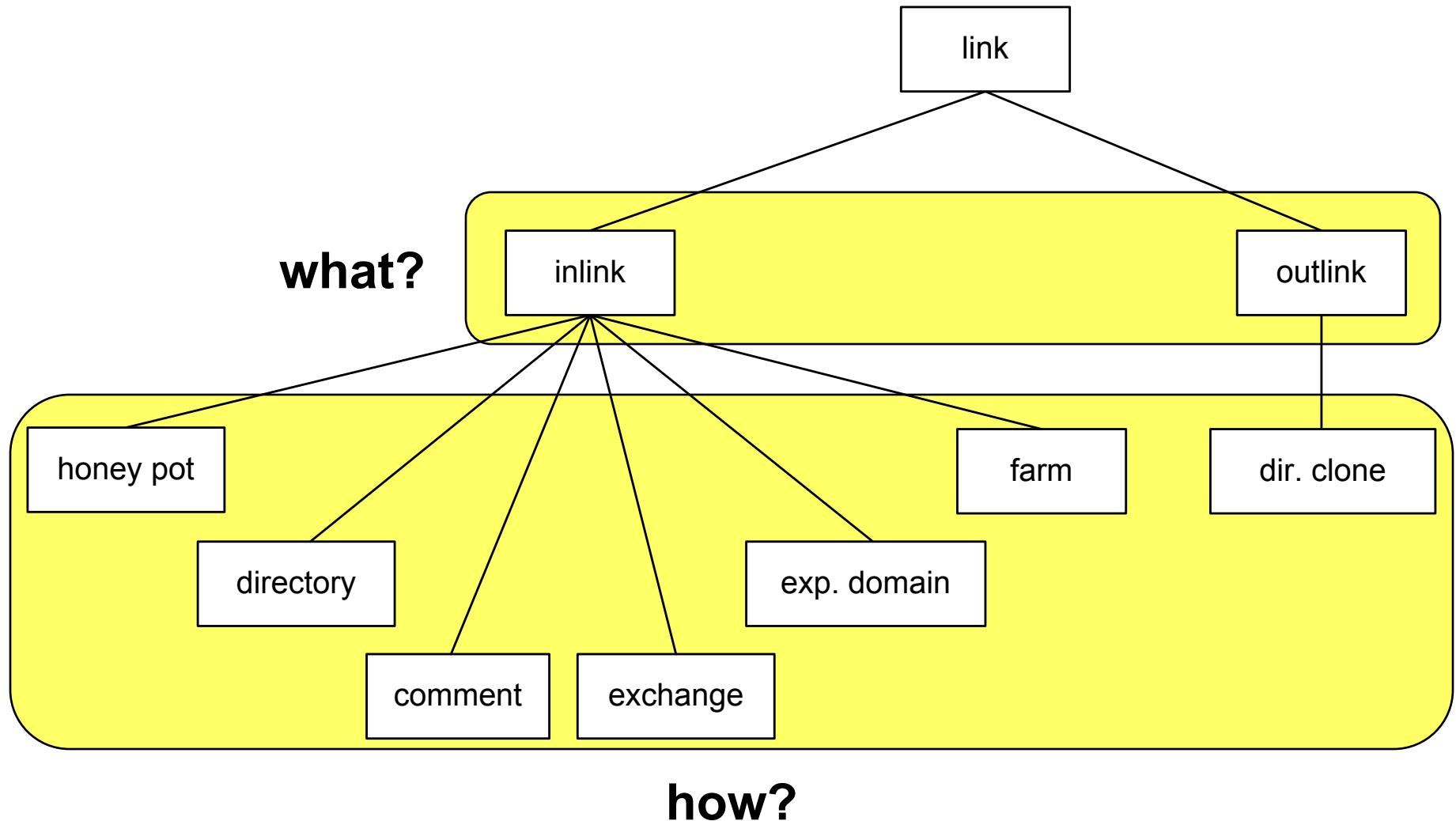


what?

Techniques / Boosting / Term

- repetition repetition **repetition**
repetition repetition repetition
- dumortierite dumose dumous dump
dumpage dumper dumpily dumpiness
dumping dumpish dumpishly
- *work in **weaving** three-women teams* is an
ancient textile art *on looms*
- *please refrain from using the **phrase***
stitching wounds located on the lower limbs

Techniques / Boosting / Link

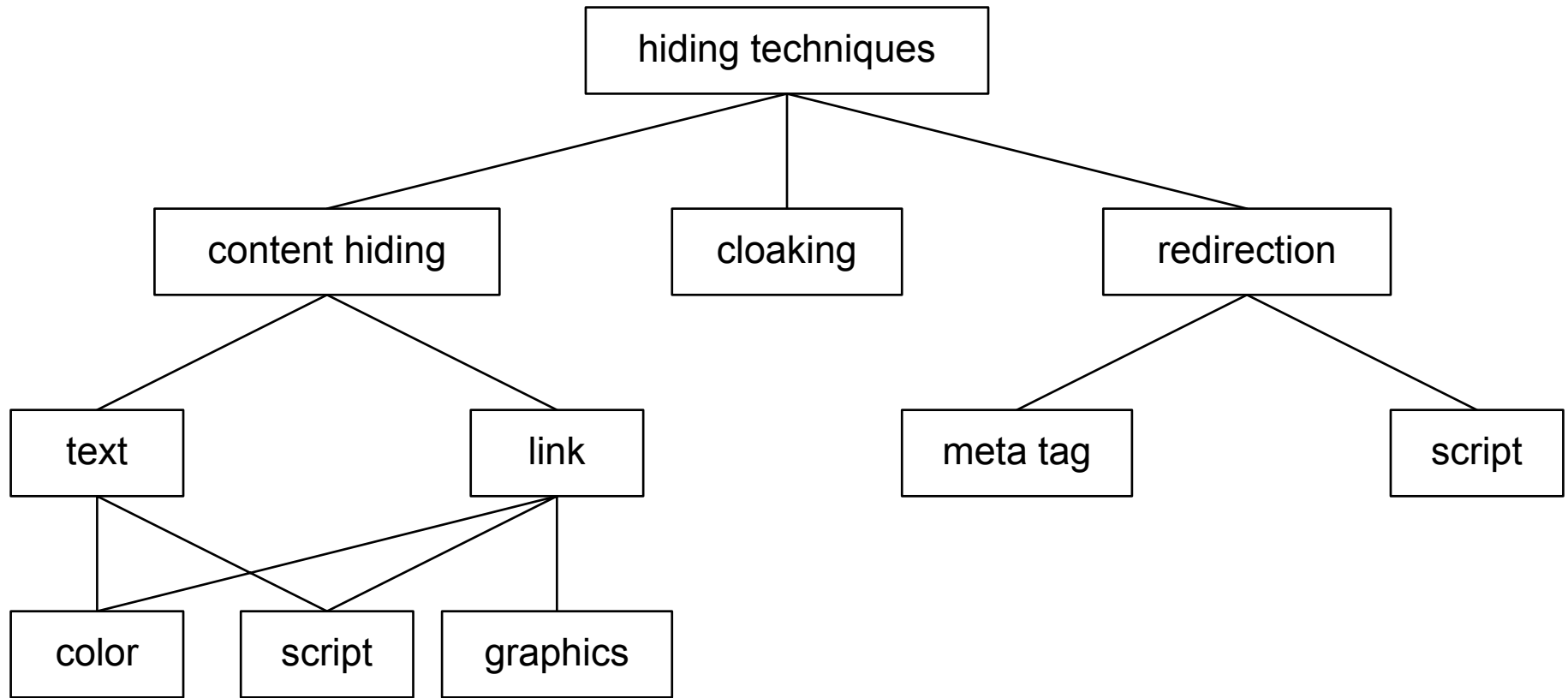


Techniques / Boosting / Link

- Directory clones
 - Duplicate (parts of) DMOZ
- Comment spam
 - Post messages (containing links) to
 - Blogs
 - (Unmoderated) forums
 - Wikis
- Link spam farms
 - Increase size
 - Increase collusion

Techniques / Hiding

- Used to conceal boosting



Techniques / Hiding

■ Content hiding

```
<style type = "text/css">  
  body {  
    background-color: white;  
    color: white; }  
</style>
```

```
<a href = "..."><img src =  
"1x1.gif"></img></a>
```

```
<div style = "visibility: hidden">You can't  
see me!</div>
```

■ Cloaking

- Identify web crawlers
- Serve a different version of the page

Roadmap

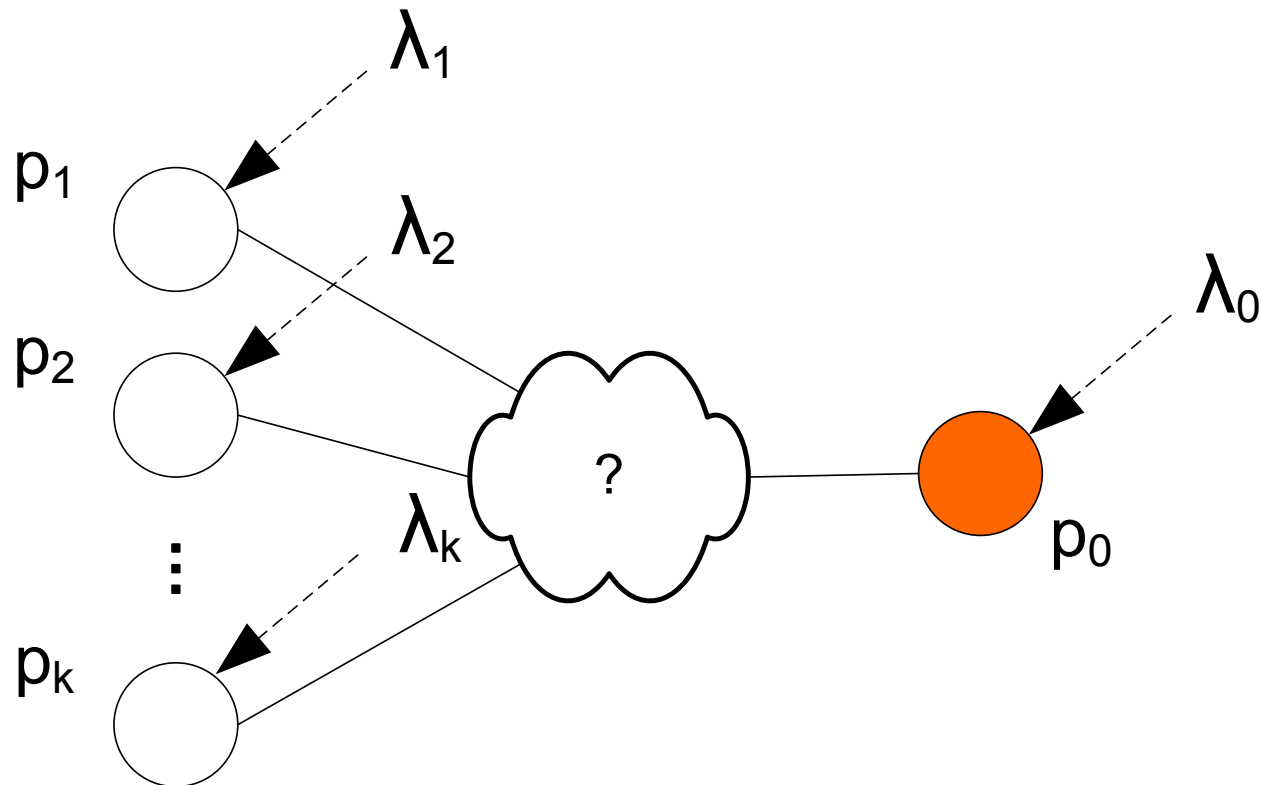
- What is search engine spam?
- What techniques do spammers use?
- Work at Stanford
- Challenges ahead

Work at Stanford

- Analysis
 - Link spam farms and alliances
- Demotion
 - TrustRank
- Detection
 - Spam mass estimation
 - See publications

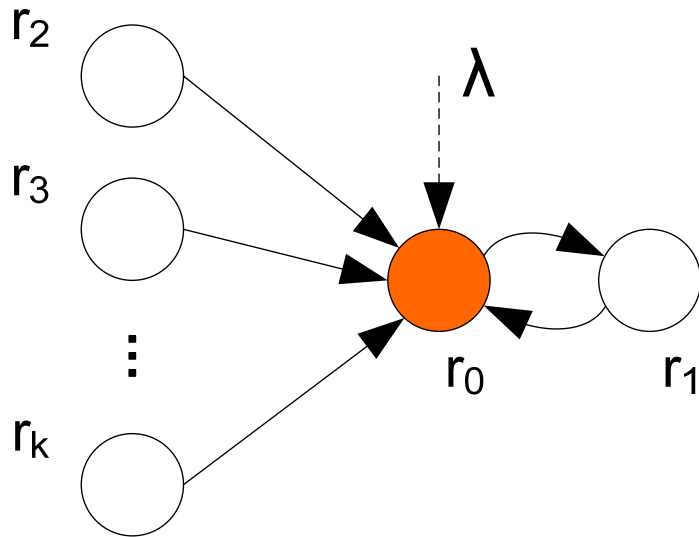
Link Spam Farms & Alliances

- Spammer's goal: increase PageRank
- Farm model



Link Spam Farms & Alliances

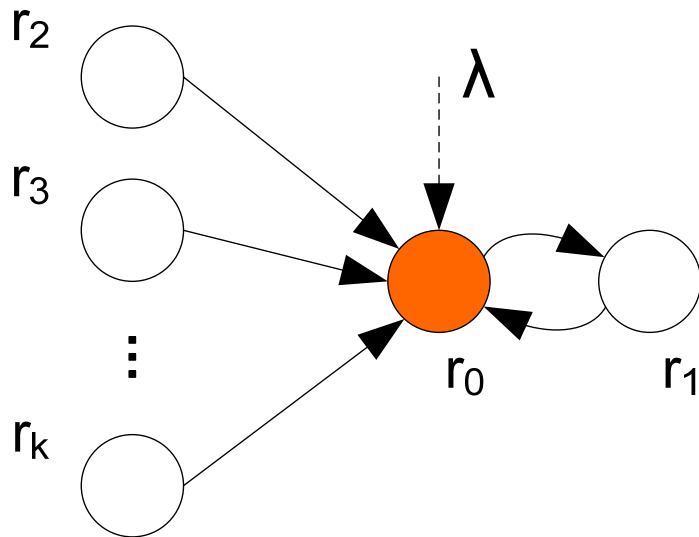
- Optimal farms
 - Short loops including target



Link Spam Farms & Alliances

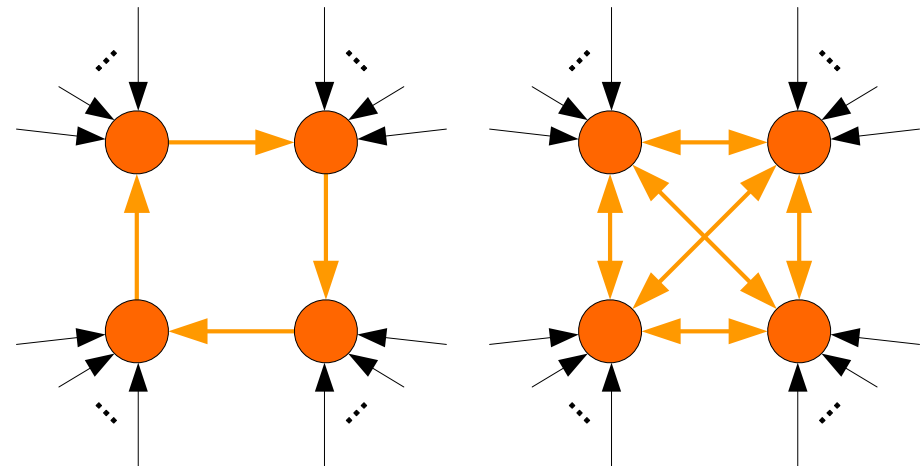
■ Optimal farms

- Short loops including target

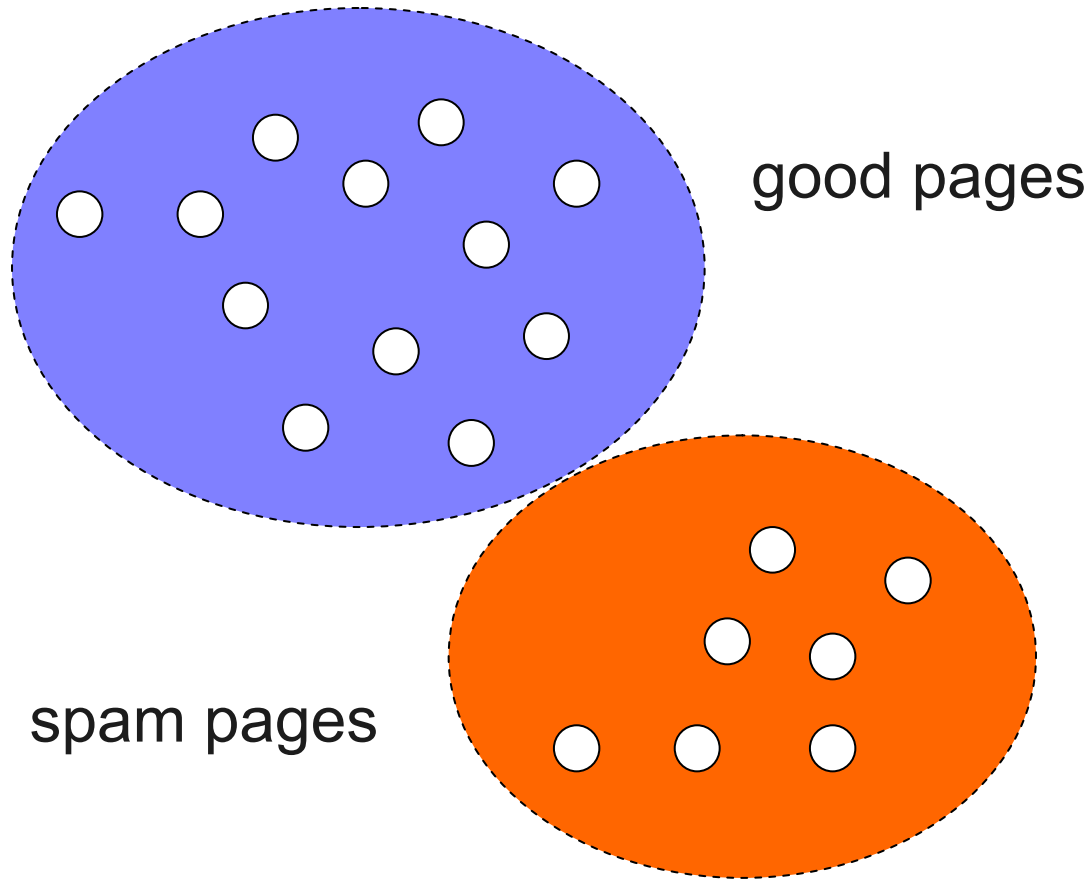


■ Alliances

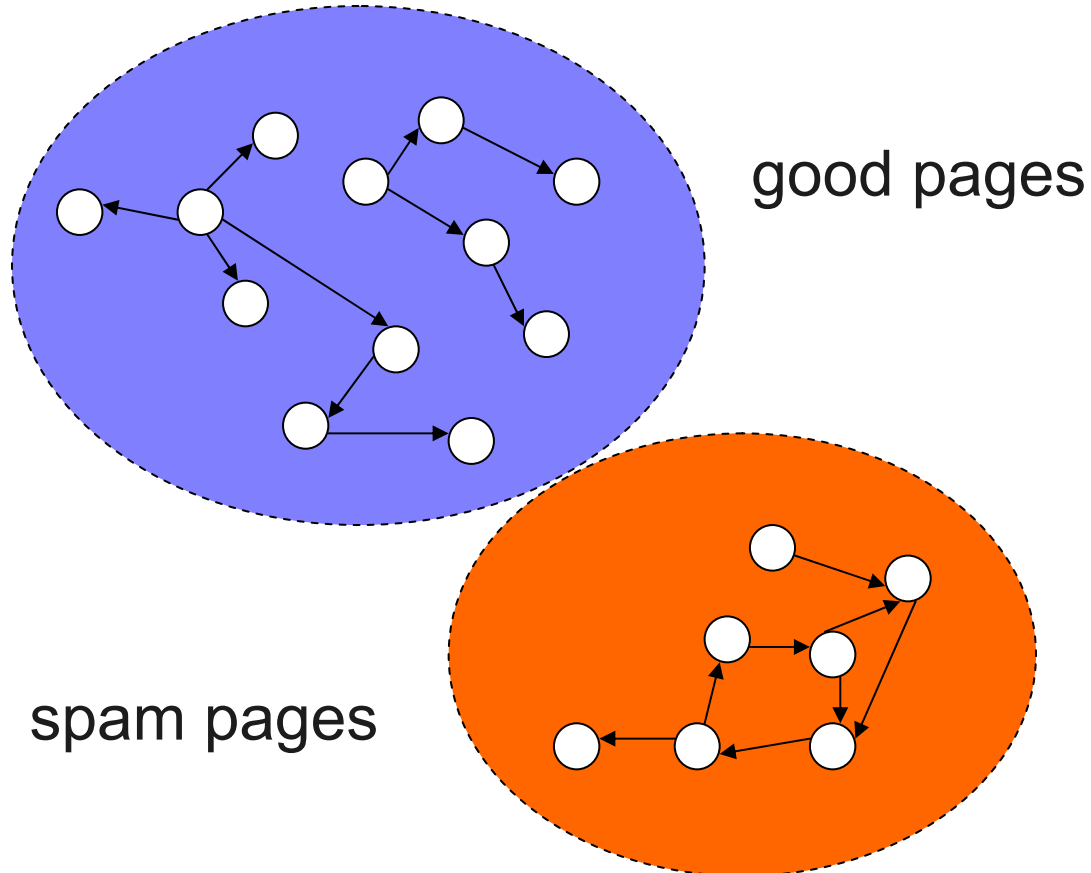
- Interconnected farms
- 2 always better than 1
- Larger alliances often benefit all



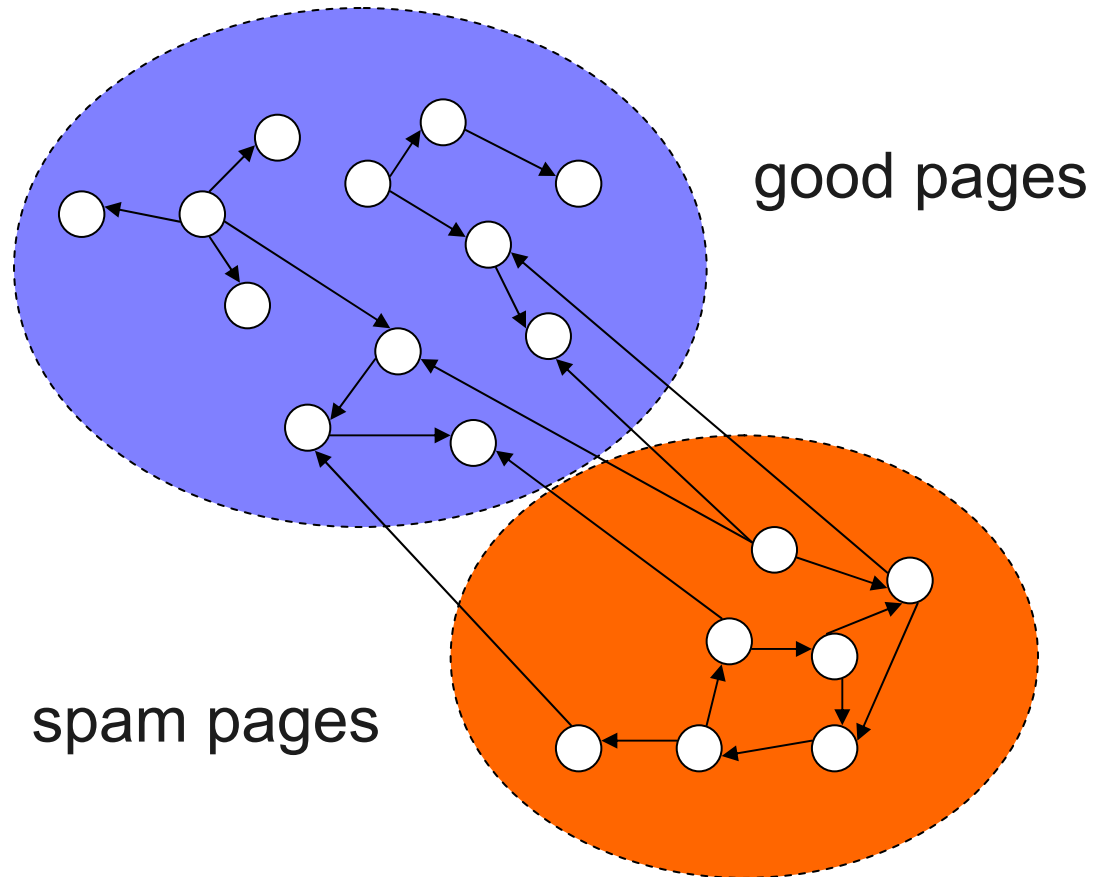
TrustRank / Observation



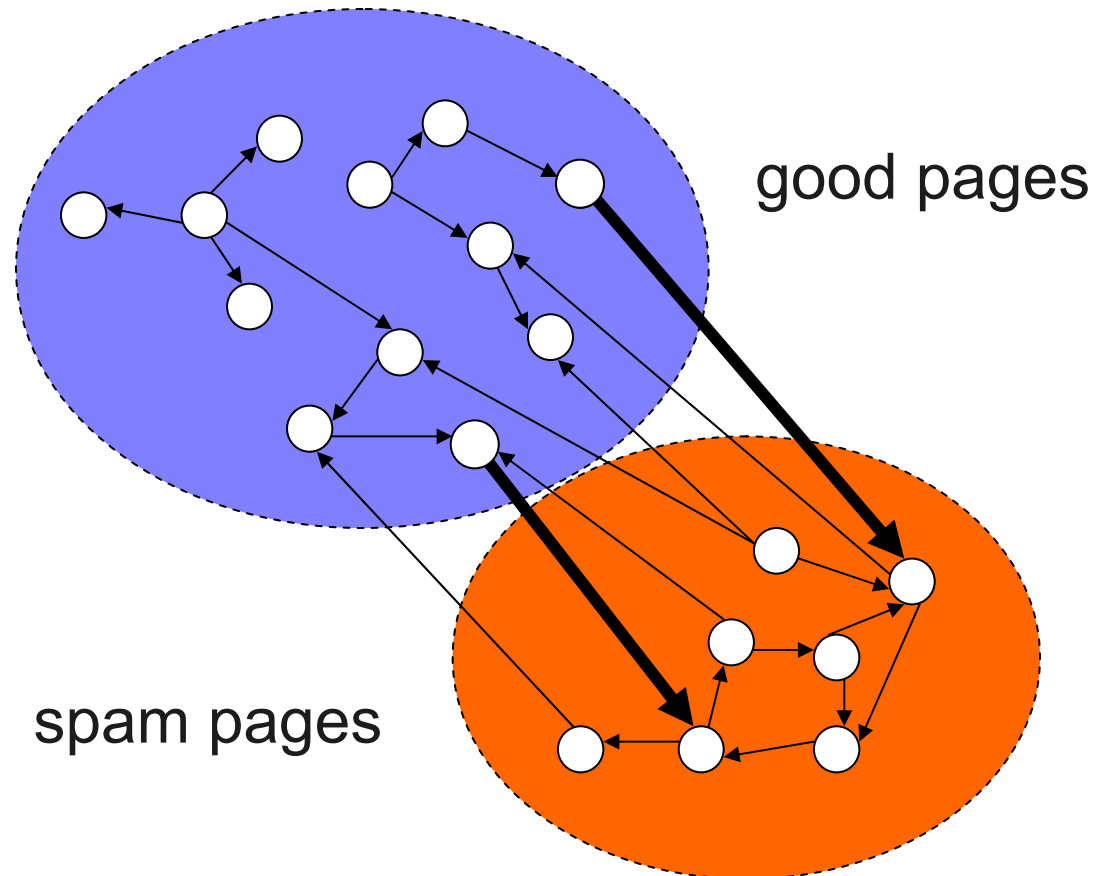
TrustRank / Observation



TrustRank / Observation



TrustRank / Observation



**Approximate isolation of good pages:
good pages seldom point to spam**

TrustRank / Objective

Separate good pages from spam pages

TrustRank / Objective

Separate good pages from spam pages

What?

Assign high scores to very good pages

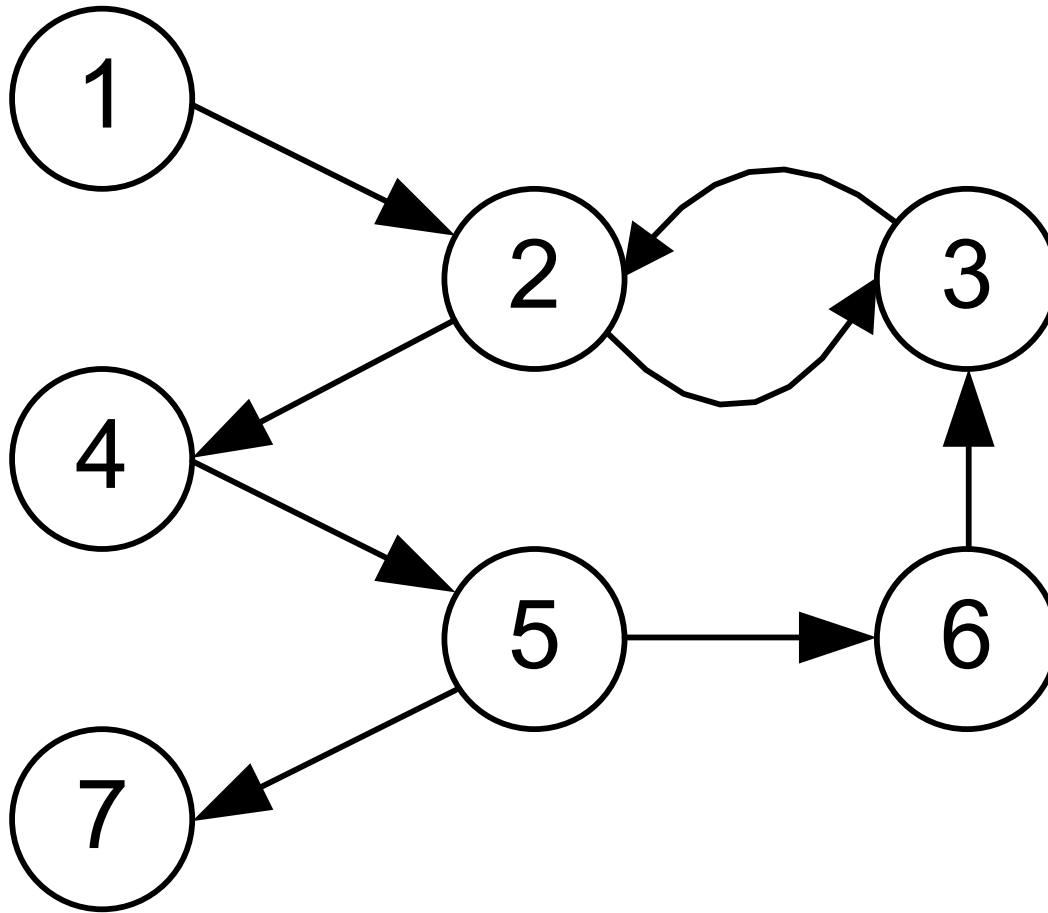
How?

Propagate scores from **known** good pages
(seed set)

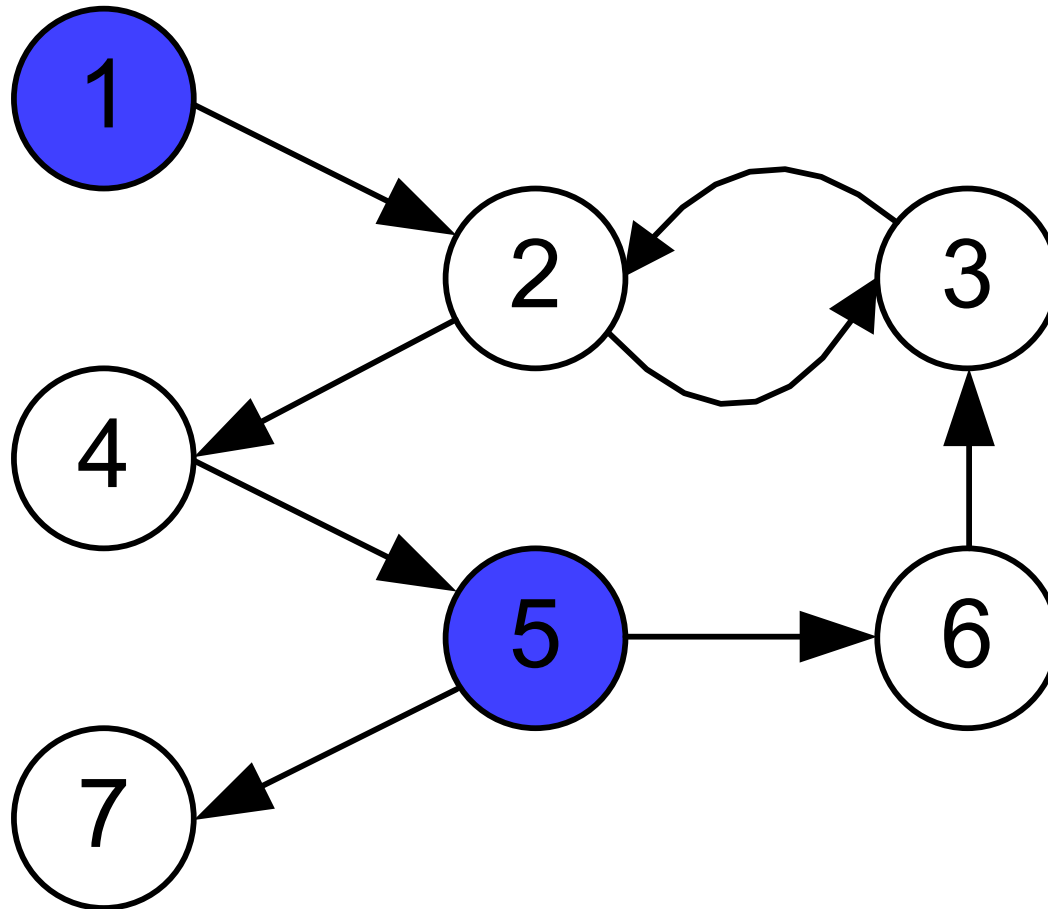
When?

Use results in ranking

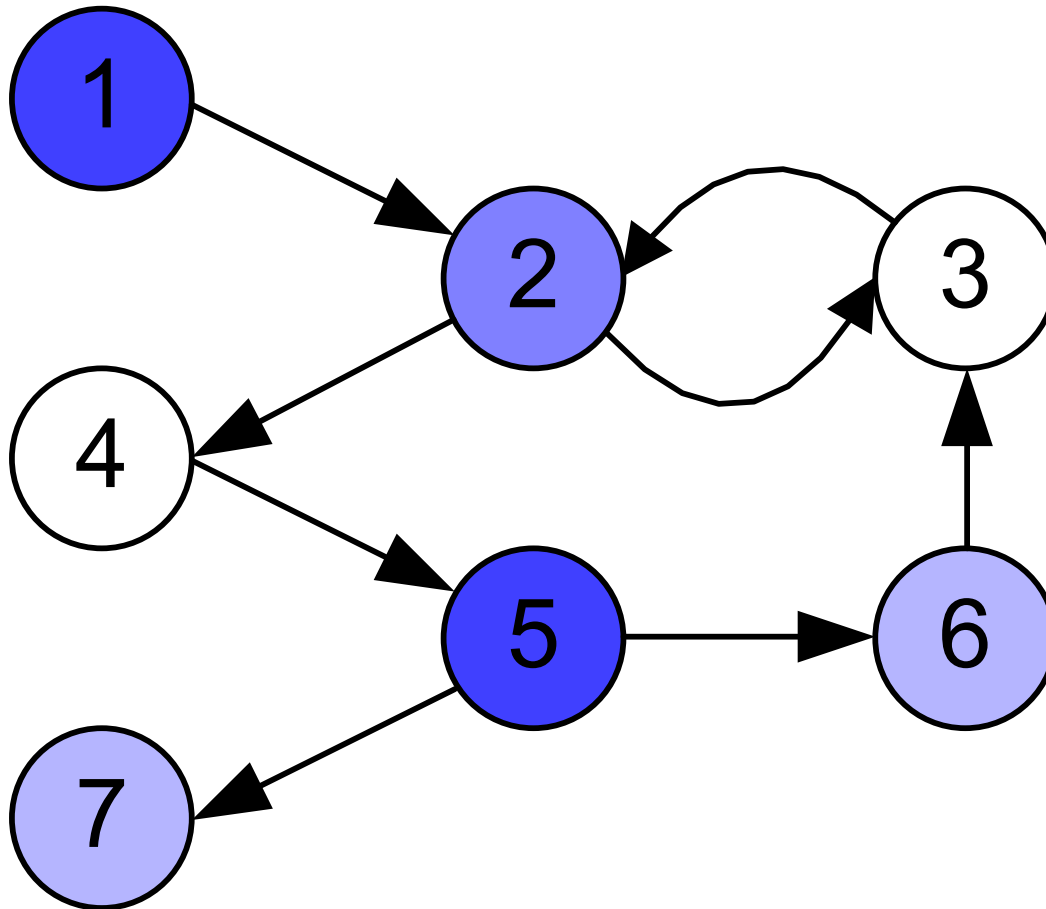
TrustRank / Example



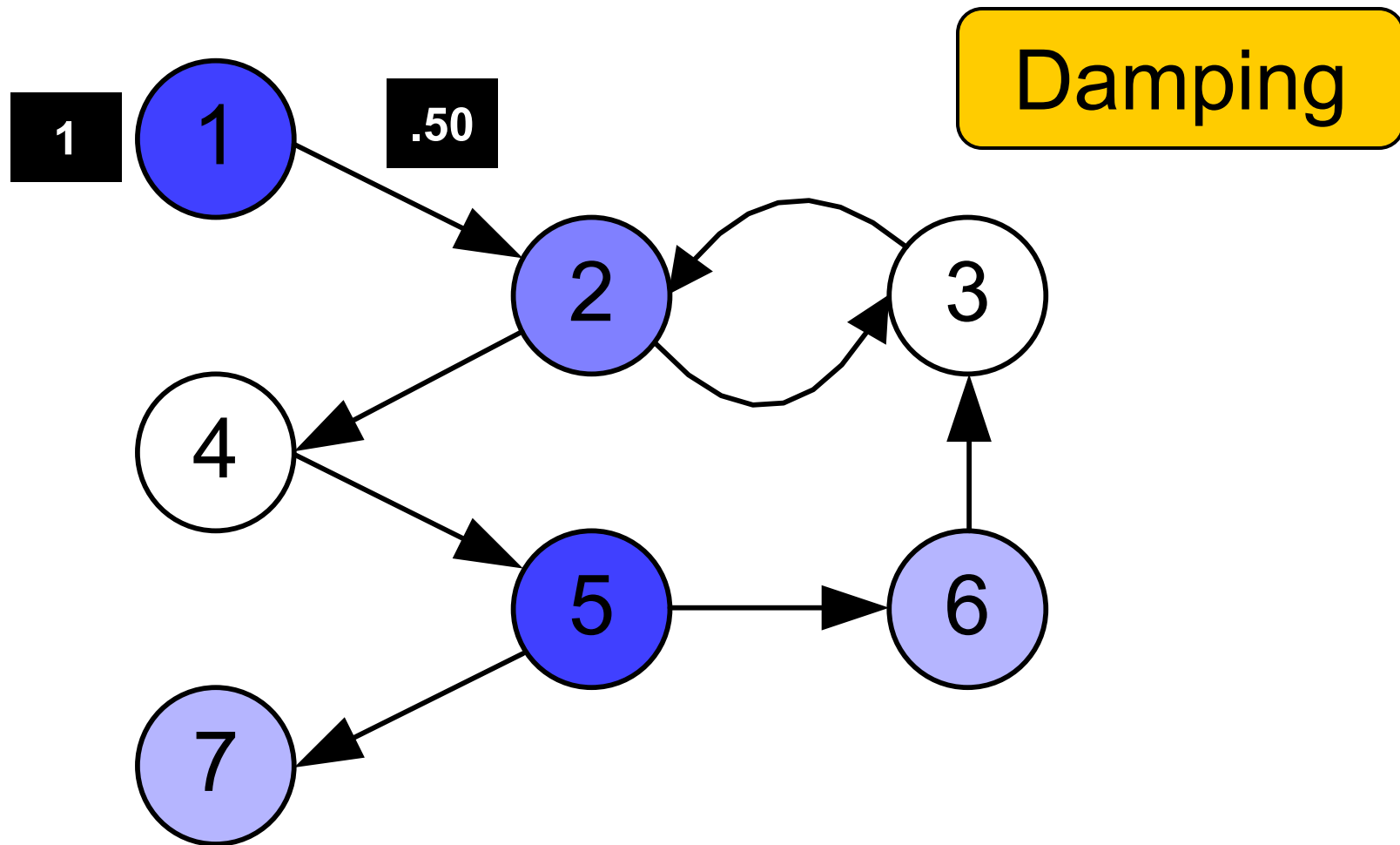
TrustRank / Example



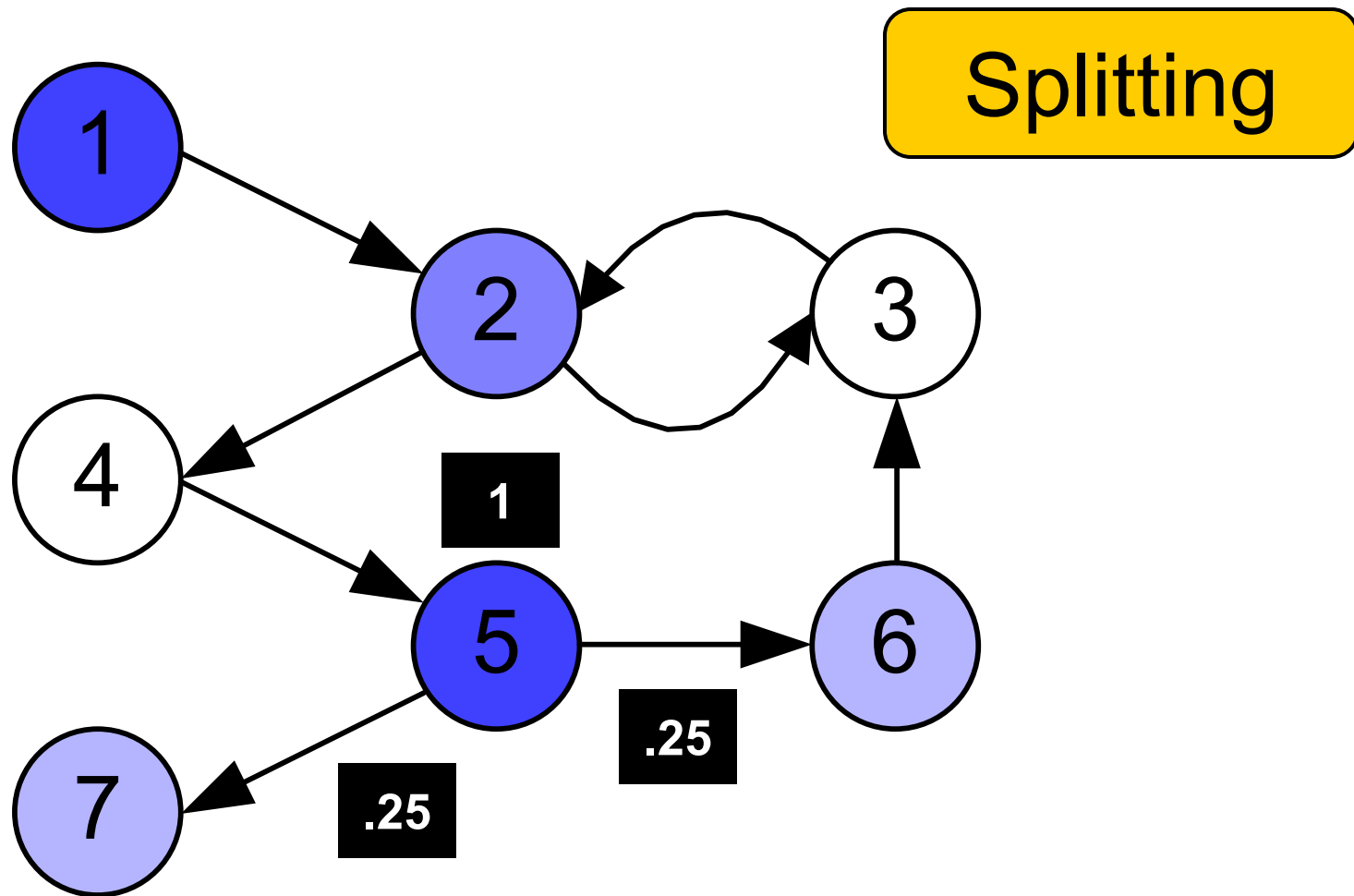
TrustRank / Example



TrustRank / Example



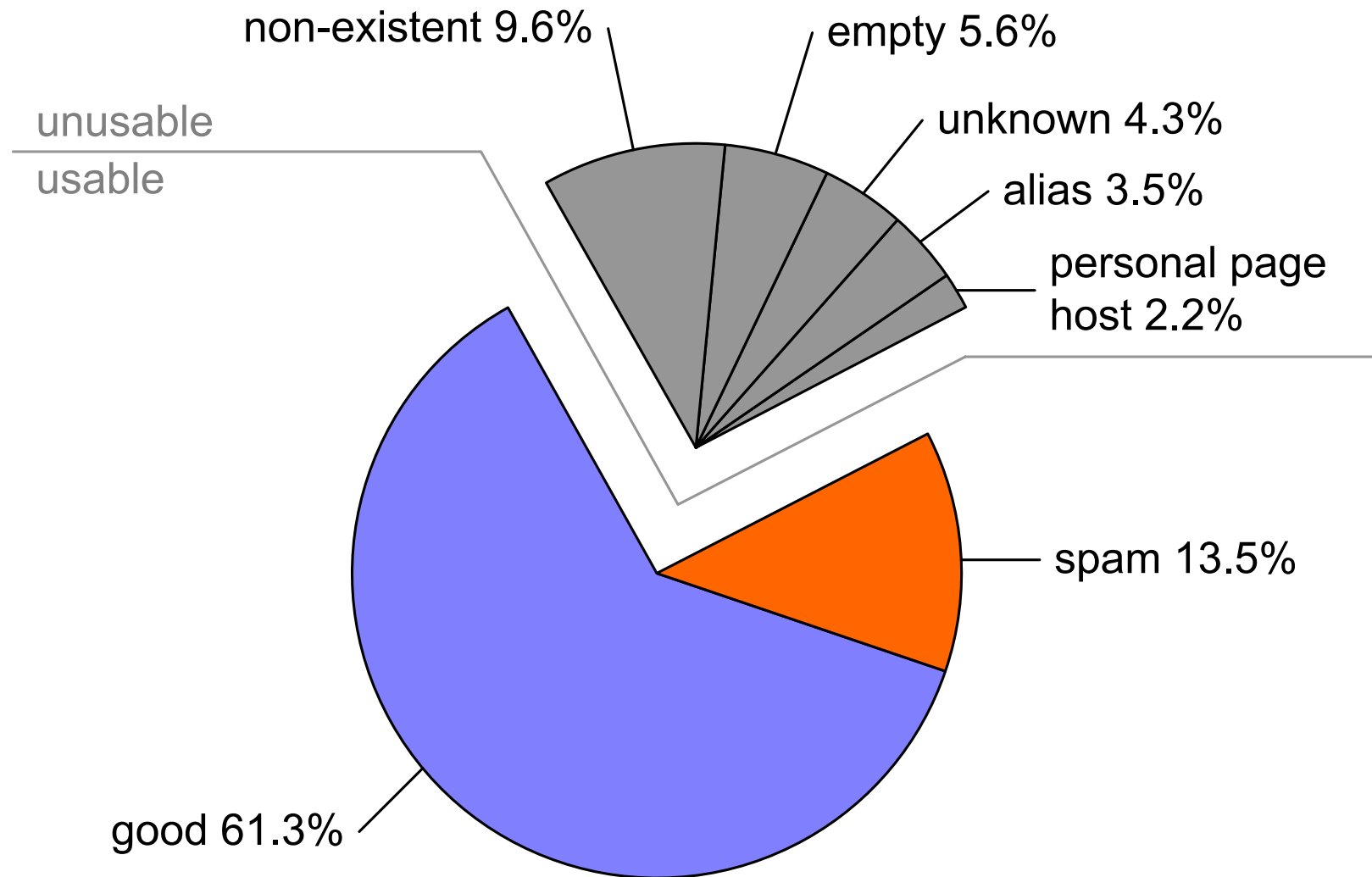
TrustRank / Example



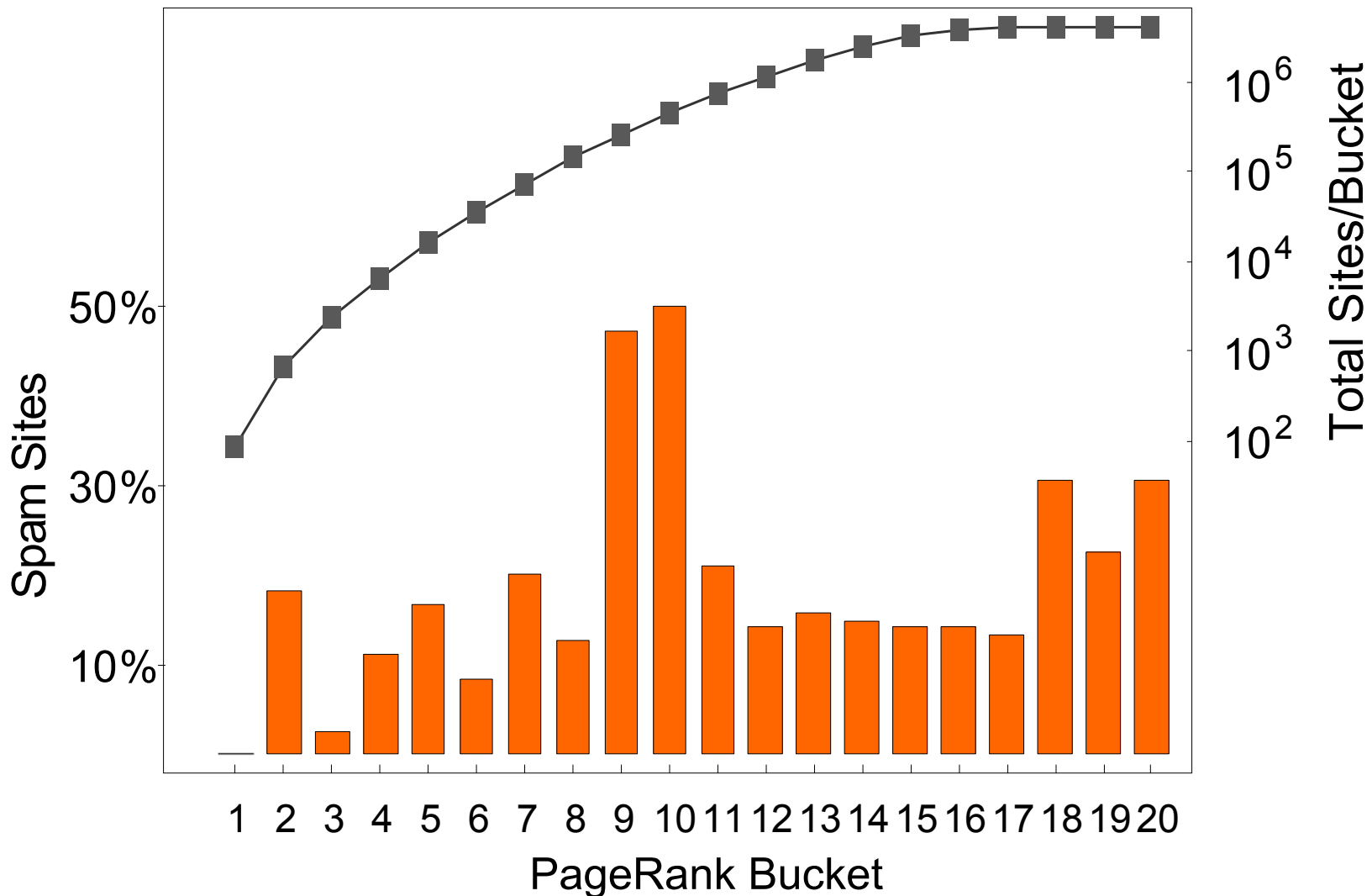
TrustRank / Experiments

- Web data
 - Entire AltaVista index (June 2003)
 - Site-level web graph
 - 31M nodes
 - 13M without inlinks
- Seed set
 - 2500 candidates
 - 178 selected high-quality sites
- Evaluation sample
 - 1000 manually tagged sites

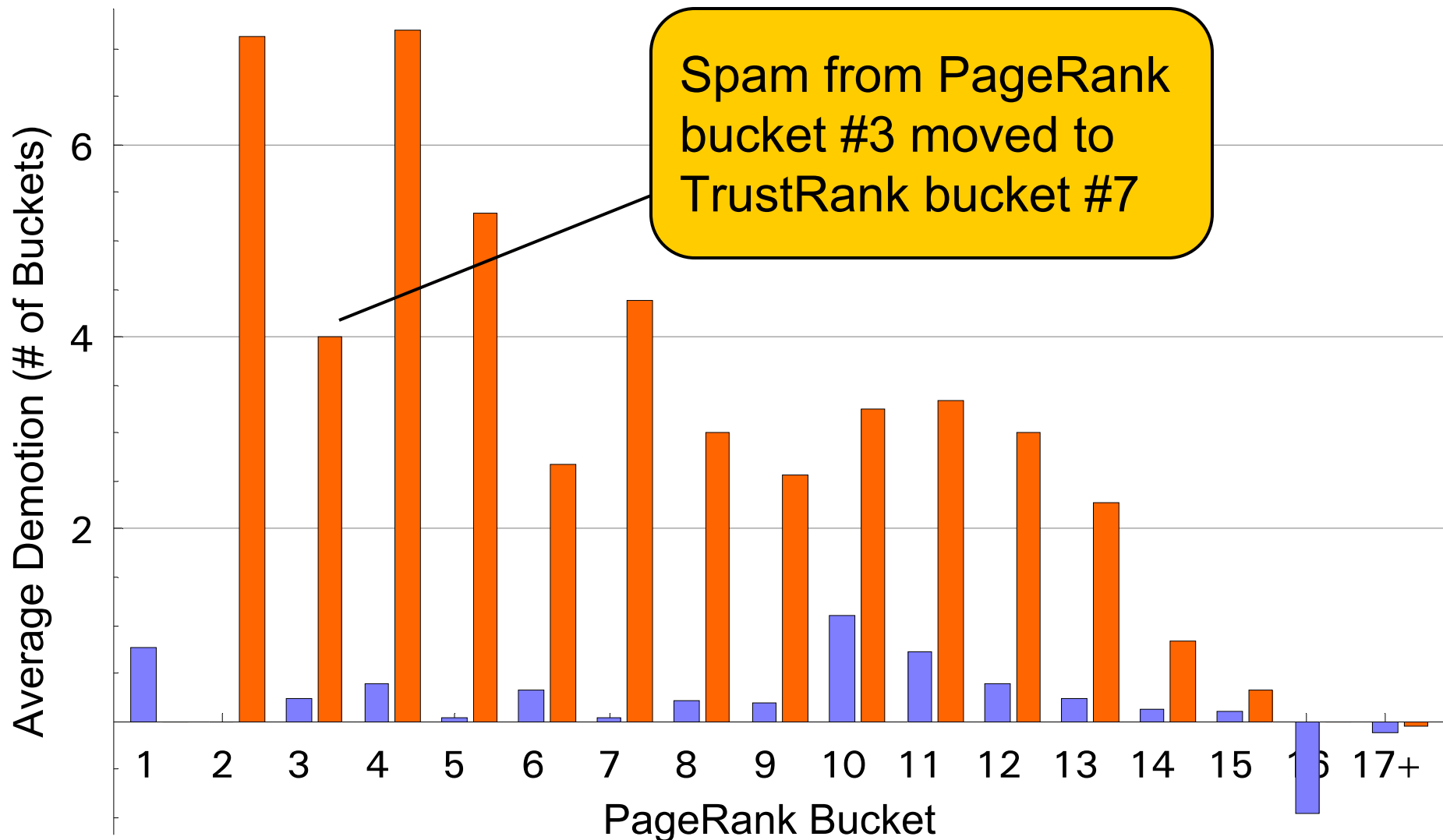
TrustRank / Experiments



TrustRank / Experiments



TrustRank / Experiments



Review

- Spamming techniques
 - Boosting
 - Term-based
 - Link-based
 - Hiding
- Techniques for combating spam

Challenges

- Remove economic incentive
 - Why not just charge for high ranking?
 - Revenue based on transactions generated, not click-through rate
- Spam on social networking sites
 - Flickr, MySpace, ...
- Mechanism design
 - Spam-proof algorithms/services

Challenges

- Better use of human judgment
 - Exploitation of implicit feedback
- Spam prevalence statistics
 - Evolution over time
 - At various levels of granularity
 - In search engine index vs. query results

Thank You!

- **Stanford InfoLab**
<http://infolab.stanford.edu>
- **Publications**
<http://dbpubs.stanford.edu>
- **Contact**
hector@cs.stanford.edu
zoltan@cs.stanford.edu