



Web Spam, Propaganda and Trust

P. Takis Metaxas
Computer Science Department
Wellesley College

Joint work with Joe DeStefano

Outline of the Talk

- ◆ The Web and its Spam
 - ◆ A Short History of the Search Engines
 - ◆ Web Spam as Propaganda ...
 - Propaganda Primer
 - ◆ Anti-propagandistic techniques on Spam
 - Experimental Results
- ◆ Conclusions and Next Steps ..

The Web ...

- ◆ Has changed the way we get informed
- ◆ Has changed the way we make decisions
(financial, medical, political, ...)
- ◆ Is huge
 - 2-10 billion static pages publicly available,
 - ◆ doubling every year
 - Three times this, if you count the “deep web”
 - Infinite, if you count dynamically created pages
- ◆ Will be omnipresent
 - Computers, Cell phones, PDA’s, thermostats, toasters ...
- ◆ Can be unreliable

JUMP START by Robb Armstrong



... and its Spam



Web [Images](#) [Groups](#) ^{New!} [News](#) [Froogle](#) [more »](#)

can human growth hormone increase muscle mass?

Search

[Advanced Search](#)
[Preferences](#)

Web

Results 1 - 10 of about 313,000 for [can human growth hormone increase muscle mass?](#). (0.26 seconds)

[Human Growth Hormone](#)

www.HGH-Info-Buyers-Guide.org

Doctor Recommendations & Warnings. Find Info on Effective & Safe HGH!

Sponsored Link

Tip: Have a question? Ask the researchers at [Google Answers](#).

[Human Growth Hormone Natural HGH Supplement Decrease Fat and Gain ...](#)

... **Increase** stamina. Improve immune functioning. ... Now, you **can** benefit from the premiere HGH ... information you need on hGH, **human growth hormone** hGH releaser, youth ...
www.hghplus.net/ - 23k - [Cached](#) - [Similar pages](#)

[Human Growth Hormone - "Rejuvenation" Medicine?](#)

... the list of the age-depleted **hormones** that **can** be replaced ... is natural; it is plentiful in the **human** body during ... weight gain of 30-50 pounds occur when HGH wanes ...
www.usdoctor.com/gh.htm - 17k - [Cached](#) - [Similar pages](#)

[Secretagogue-One can increase - hgh, hGH - Human Growth Hormone ...](#)

Maximum **Human** Performance (MHP ... almost embarrassed to carry Secretagogue-One, **growth hormone** and IGF ... Enhances all other **hormone** activities; **Increases** hGH and IGF1; ...
www.powersupplements.com/hgh.htm - 12k - [Cached](#) - [Similar pages](#)

[Best HGH for Bodybuilding - HGH Supplement](#)

... Yes, You **Can** Now Get Actual **Human Growth Hormone** ... **Human Growth Hormone** (HGH) is one of the most powerfully effective anabolic **hormones** available. ...
www.trulyhuge.com/hgh.htm - 14k - Jan 29, 2005 - [Cached](#) - [Similar pages](#)

[Human Growth Hormone HGH - Pro HGH Supplement | Human Growth ...](#)

... HGH) over the counter, however, you **can** improve your ... The same study found that administering **human growth hormone** (hgh) resulted in a 8.8% **increase** in lean ...
www.bodybuildingforyou.com/pro-hormones/human-growth-hormone-hgh.htm - 33k - [Cached](#) - [Similar pages](#)

[Buy HGH human growth hormone supplement, oral HGH hormone](#)

Somatrop Rx is a legal HGH **hormone** supplement with the strongest potency allowed by law. ... Now you **can** get actual **human growth hormones** legally without ...
www.anabolicmall.com/hgh.htm - 19k - [Cached](#) - [Similar pages](#)

Sponsored Links

[Human Growth Hormone](#)

US Prescribed HGH and Testosterone
Amazing results. Great service.
www.AntiAgingGroup.com

[Human Growth Hormones](#)

Don't buy HGH until you read this
Get the latest research and reviews
www.hghreview.net

[Muscle Without Steroids](#)

Step by step method to pack on
new **muscle** every workout.
www.PrecisionTraining.com

[HGH - The Truth](#)

What They Don't Want You To Know
What It Does To Your Body Affiliate
www.GenF20.com

[Muscle Growth](#)

Want Faster **Muscle Growth**?
Learn The Tips & Secrets
www.steroid-encyclopaedia.com

[Build muscle fast](#)

System used by 150,000+ worldwide
Put on 30+ pounds of pure **muscle**!
www.BestWeightliftingProgram.com

[Muscle Growth Secrets](#)

... and its Spam



Web [Images](#) [Groups](#) ^{New!} [News](#) [Froogle](#) [more »](#)

is adhd a real disease?

Search

[Advanced Search](#)
[Preferences](#)

The following words are very common and were not included in your search: is a. [\[details\]](#)

Web

Results 1 - 10 of about 200,000 for [is adhd a real disease?](#). (0.18 seconds)

Tip: Have a question? Ask the researchers at [Google Answers](#).

[Texans for Safe Education Resolution](#)

Texans for Safe Education Is **ADHD A Real Disease?** Dr. Fred Baughman is a neurologist who has discovered **real** diseases. By directly ...

www.wildestcolts.com/safeEducation/real.html - 8k - [Cached](#) - [Similar pages](#)

[Truths About ADHD and Stimulant Drugs](#)

... Informative Websites. www.adhdfraud.org - Dr. Fred Baughman's excellent website, containing the best of his essays revealing that **ADHD** is not a **real disease**. ...

www.wildestcolts.com/mentalhealth/stimulants.html - 14k - [Cached](#) - [Similar pages](#)

[[More results from www.wildestcolts.com](#)]

[To all website visitors: I urge one and all concerned with the ...](#)

MAKE CHADD ANSWER: IS **ADHD A REAL DISEASE**—YES OR NO? by Fred A. Baughman Jr., MD, May 26, 2001. Is **ADHD a real disease**—yes or no? ...

www.adhdfraud.org/commentary/5-27-01-1.htm - 5k - [Cached](#) - [Similar pages](#)

[ADD and ADHD Fraud. Find out the truth about ADD and ADHD. ...](#)

... the fraud of Attention Deficit Hyperactivity Disorder (**ADHD**)—Compiled by ... Making "**disease**" (**real** diseases—epilepsy, brain tumor, multiple sclerosis, etc.) or ...

www.adhdfraud.org/ - 21k - Jan 29, 2005 - [Cached](#) - [Similar pages](#)

[[More results from www.adhdfraud.org](#)]

[ADHD a Real Disorder](#)

... Some people say, **ADHD** is not a **real** disorder because ... He says : "**ADHD** is a disorder that cannot be identified in the same way as polio, heart **disease** or other ...

web4health.info/en/answers/adhd-real-disorder.htm - 20k - [Cached](#) - [Similar pages](#)

[Baughman Dispels The Myth of ADHD](#)

... The "**disease**," Baughman tellsInsight, "is a total 100 percent ... to bring an end to the **ADHD** diagnosis ... as an adult and child neurologist, diagnosing **real** diseases. ...

www.becomehealthynow.com/article/dirty/209 - 62k - [Cached](#) - [Similar pages](#)

What is Web Spam?

- ◆ The practice of manipulating web pages in order to cause search engines rank them higher than they would without manipulation
- ◆ “...than they deserve”
- ◆ “... unjustifiably favorable [ranking wrt] the page’s true value”
- ◆ “...unethical web page positioning”
- ◆ It is a problem, not only for search engines
 - Primarily for users
 - As well as for content providers
- ◆ It is first a **social** problem, then a technical one

Who is Spamming and Why?

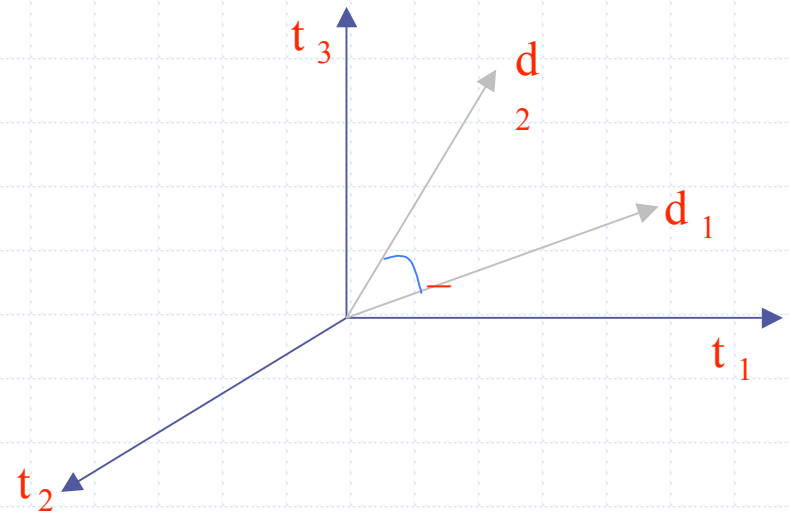
- ◆ Companies
 - Big companies
 - Small businesses
 - ◆ Advertisers and Promoters
 - Search Engine Optimizers
 - ◆ Special interest groups
 - Religious interests
 - Financial interests
 - Medical interests
 - Political interests
 - etc
 - ◆ Everybody could/would
 - My doctor
 - You (?), Me (!)
- ◆ 85% of searchers do not go beyond top-10
 - ◆ People (still) trust the written word
 - ◆ People trust the search engines

A Short History of Search Engines

- ◆ 1st Generation (ca 1994):
 - AltaVista, Excite, Infoseek...
 - Ranking based on **Content**
 - ◆ Pure Information Retrieval
- ◆ 2nd Generation (ca 1996):
 - Lycos
 - Ranking based on **Content + Structure**
 - ◆ Site Popularity
- ◆ 3rd Generation (ca 1998):
 - Google, Teoma
 - Ranking based on **Content + Structure + Value**
 - ◆ Page Reputation
- ◆ In the Works
 - Ranking based on “the need behind the query”
 - ◆ ??

1st Generation: Content Similarity

- ◆ Boolean operations on query terms did not go very far
- ◆ *Content Similarity Ranking:*
*The more **rare** words two documents share, the more similar they are*
- ◆ Similarity is measured by vector angles
- ◆ Query Results are **ranked** by sorting the angles between query and documents
- ◆ **How To Spam?**



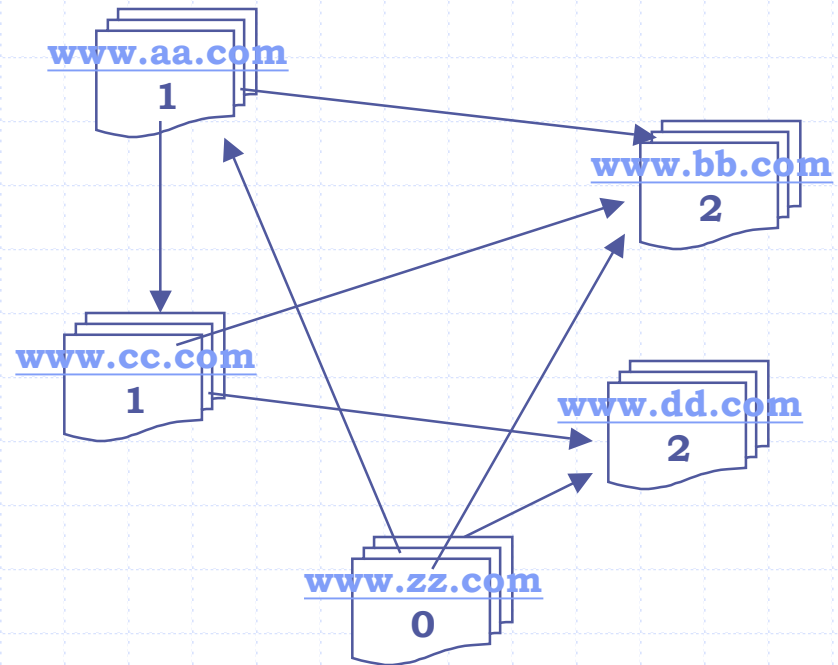
1st Generation: How to Spam

- ◆ Add keywords so as to confuse page relevance
- ◆ Hide them from human eyes
- ◆ Searching for Jennifer Aniston?

SEX SEXY MONICA LEWINSKY JENNIFER LOPEZ CLAUDIA SCHIFFER CINDY CRAWFORD
JENNIFER ANNISTON GILLIAN ANDERSON MADONNA NIKI TAYLOR ELLE MACPHERSON KATE
MOSS CAROL ALT TYRA BANKS FREDERIQUE KATHY IRELAND PAM ANDERSON KAREN MULDER
VALERIA MAZZA SHALOM HARLOW AMBER VALLETTA LAETITA CASTA BETTIE PAGE HEIDI
KLUM PATRICIA FORD DAISY FUENTES KELLY BROOK SEX SEXY MONICA LEWINSKY JENNIFER
LOPEZ CLAUDIA SCHIFFER CINDY CRAWFORD **JENNIFER ANNISTON** GILLIAN ANDERSON
MADONNA NIKI TAYLOR ELLE MACPHERSON KATE MOSS CAROL ALT TYRA BANKS FREDERIQUE
KATHY IRELAND PAM ANDERSON KAREN MULDER VALERIA MAZZA SHALOM HARLOW AMBER
VALLETTA LAETITA CASTA BETTIE PAGE HEIDI KLUM PATRICIA FORD DAISY FUENTES
KELLY BROOK SEX SEXY MONICA LEWINSKY JENNIFER LOPEZ CLAUDIA SCHIFFER CINDY
CRAWFORD **JENNIFER ANNISTON** GILLIAN ANDERSON MADONNA NIKI TAYLOR ELLE
MACPHERSON KATE MOSS CAROL ALT TYRA BANKS FREDERIQUE KATHY IRELAND PAM
ANDERSON KAREN MULDER VALERIA MAZZA SHALOM HARLOW AMBER VALLETTA LAETITA CASTA
BETTIE PAGE HEIDI KLUM PATRICIA FORD DAISY FUENTES KELLY BROOK SEX SEXY MONICA
LEWINSKY JENNIFER LOPEZ CLAUDIA SCHIFFER CINDY CRAWFORD **JENNIFER ANNISTON**
GILLIAN ANDERSON MADONNA NIKI TAYLOR ELLE MACPHERSON KATE MOSS CAROL ALT TYRA
BANKS FREDERIQUE KATHY IRELAND PAM ANDERSON KAREN MULDER VALERIA MAZZA SHALOM
HARLOW AMBER VALLETTA LAETITA CASTA BETTIE PAGE HEIDI KLUM PATRICIA FORD DAISY
FUENTES KELLY BROOK

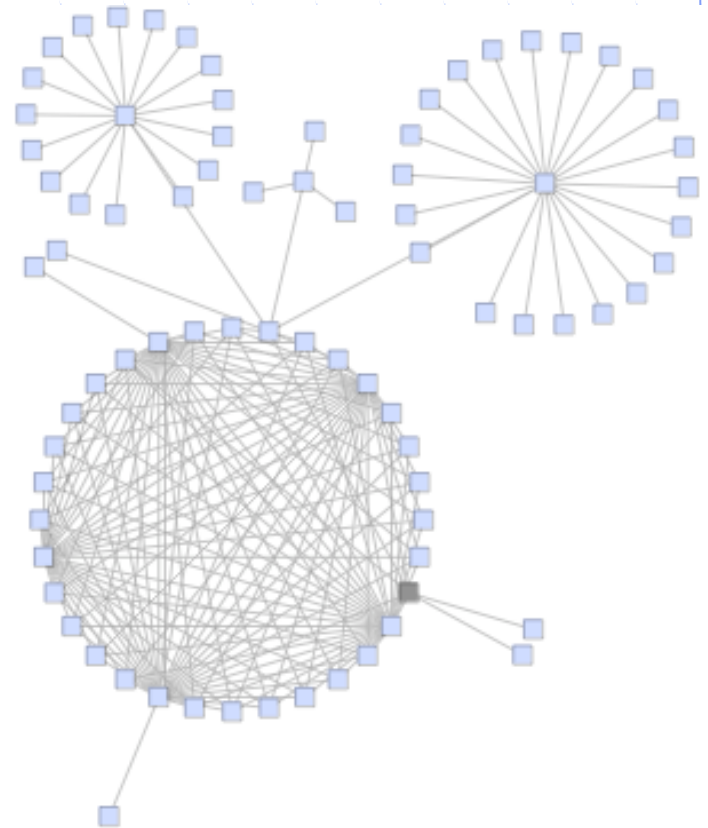
2nd Generation: Site Popularity

- ◆ A link from a page in site A to some page in site B is considered a **popularity** vote from A to B
- ◆ Rank similar pages according to popularity
- ◆ Related implementation of Popularity: DirectHit's Click-throughs
- ◆ Rich get richer: users will always try first few links returned
- ◆ **How To Spam?**



2nd Generation: How to Spam

- ◆ Heavily interconnected
“link farms”
spam popularity
- ◆ Clicking robots
spam click-throughs



3rd Generation: Page Reputation

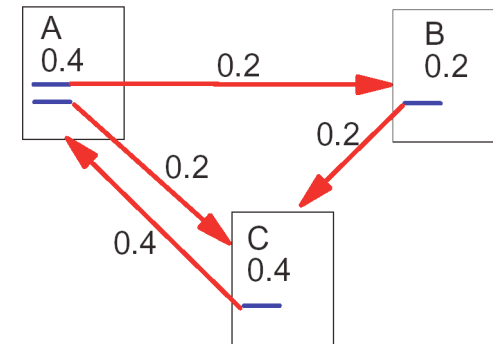
- ◆ A link from a page P_x to page P_y is considered a **confidence** vote from P_x to P_y
 - Confidence builds **reputation** (as in academic co-citations)

$$P_i = (1 - d) + d \cdot \sum_{j \rightarrow i} \frac{P_j}{C_j}$$

- ◆ The reputation “PageRank” of a page P_i = the sum of a fraction of the reputations of all pages P_j that point to P_i

- ◆ Beautiful Math behind it
 - PR = principal eigenvector of the web’s link matrix
 - PR equivalent to the chance of randomly surfing to the page
- ◆ HITS algorithm tries to recognize “authorities” and “hubs”

- ◆ **How To Spam?**



Simplified PageRank Calculation

3rd Generation: How to Spam

- ◆ Organize “mutual admiration societies” of irrelevant reputable sites

Resource Partner Additions

Currently we are only adding websites to this Resource page with PR6 or higher AND WHOSE LINKING PAGE IS PR5 OR HIGHER. The first thing we check when we receive a link request is the PageRank. If the page you are planning to put our link on does not have a PR5 or higher we delete the request without a response. Please honor this request. Those websites that meet the above requirements are added within a few days.

To add your website to our Directory simply do 3 things:

1) Copy this code to your website:

<!-- Start Copying here -->

```
<P><A HREF="http://www.1st-Hgh.com" TARGET="_blank"><B>1st Hgh, Homeopathic Human Growth Hormone Spray</B></A><B> </B>- The primary hormone in the body, information and sales. Hgh is the ultimate youth reviving hormone. For everyone over 30 years old. This product is safe, inexpensive, and comes with a full money back guarantee.</P>
```

<!-- End Copying here -->

It should look like this:

[1st Hgh, Homeopathic Human Growth Hormone Spray](http://www.1st-Hgh.com) - The primary hormone in the body, information and sales. Hgh is the ultimate youth reviving hormone. For everyone over 30 years old. This product is safe, inexpensive, and comes with a full money back guarantee.

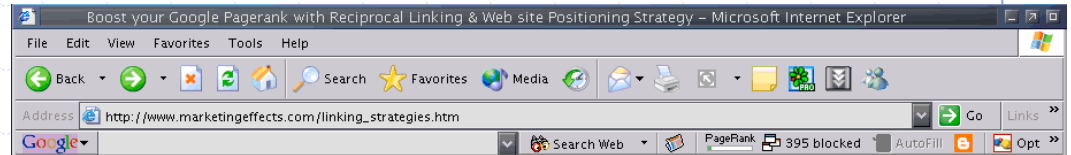
2) [E-Mail us](#) the url where we can view our link on your website (it must be PR4 or higher!);

3) Include in this e-mail your website url, title, and description.

Thank You, Partner!

An Industry is Born

- ◆ “SE Optimizer” Companies
- ◆ Advertisement Consultants
- ◆ Conferences



High PR Link Club

http://www.highprlinkclub.com/

Later... Fun News MMedia CS Funds Conf Students CriThi e-Spam Sports Hours > April

Besides, don't you have better things to do with your time than chase down links? Like run your business and make money?

Yet not only is this club a huge time saver... **many of these link pages are undiscovered little gems**... meaning you can gain a whole bunch of PageRank from them as you won't be sharing the PR with hundreds of other sites.

And get this...

There are even some PR6 pages I've found that hardly anyone knows about and they're more than willing to swap links with any decent website!

Are you starting to see the big picture here?

As if all that weren't enough... what if I also made it so simple for you to conduct your link campaign that it's as easy as

copy, paste and forget it...

But wait... here's the best part...

... what if I showed you how to do the whole thing in

... about 1 hour a month!

Yes, you read that right... **about one hour a month is all it will take you to boost your pagerank, increase your search engine rankings and easily grab more traffic!**

Now if you really do 'get' what I'm offering, you're practically ready to knock me over to get into the private, Members Only section of this Website, right?

(Just think... this entire, awesome program is that close at hand! Now I saw that! Quit trying to peak around the corner... you need a password to get in!)

Get your [lifetime membership](#) to the High PR Link Club for **ONLY \$97!**

Marketing Effects

Reciprocal Link Tools and Website Positioning Strategies

Drive Thousands of Targeted Visitors to your website and boost your Google Pagerank

Why Having Links to Your Website Are Important

Having links from other websites is very important, yet, it's one of the most overlooked strategies for increasing traffic and for increasing your position in the search engines.

As soon as you launch a website, you should spend some time working on a strategic web site linking strategy.

By working on getting sites to link to you, you'll not only see an increase in traffic, but you'll also see your website achieve better listings in the search engines, especially the most important search engine - Google.

How to find good sites to link to you

It's all about getting the right sites to link to you. Just remember **quality before quantity**. That means to avoid links farms and FFAs like the plague. Many search engines will penalize you for having a link from a link farm. Ask yourself: "Would I want people to associate me with this website?" If the answer is no, then move on to the next site. With over 40,000,000 websites, you should be able to find some high quality websites willing to link to you.

Before you get started, I recommend these two tools, as they'll make finding the quality sites you want to be linked to much easier.

The first is the [Google Toolbar](#). With the Google Toolbar, you'll be able to see how important Google thinks the page is.

The second is [Alexa Toolbar](#). With Alexa, you can see how popular the site is. Alexa ranks every site based on how often other toolbar users visit the site. It's an excellent tool, and the best part about both of these tools are that they're absolutely FREE.

The next step is to get on your favorite search engine (I recommend Google since it's

Internet

Web Spam as a major force behind Search Engines Evolution

◆ Search Engine's Action

- ◆ 1st Generation: Pure IR
 - Content
- ◆ 2nd Generation: Popularity
 - Content + Structure
- ◆ 3rd Generation: Reputation
 - Content + Structure + Value
- ◆ In the Works
 - Ranking based on "the need behind the query"

◆ Web Spammers Response

- ◆ Add keywords so as to confuse page relevance
- ◆ Create "link farms" of heavily interconnected sites
- ◆ Organize "mutual admiration societies" of irrelevant sites
- ◆ ??

Can you guess what they will do?

They will try to modify the Web Graph for their benefit

Is there a pattern on how to spam?

And Now For Something Completely Different(?)

◆ Propaganda:

- Attempt to **modify** human behavior, and thus **influence** their actions in ways **beneficial** to propagandists

◆ Theory of Propaganda

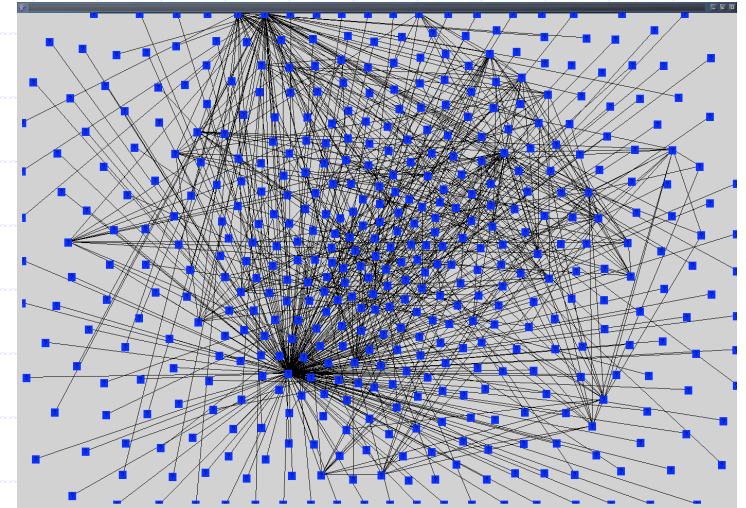
- Developed by the Institute for Propaganda Analysis 1938-1942

◆ Propagandistic Techniques (and ways of **detecting** propaganda)

- Word games
 - ◆ Name Calling
 - ◆ Glittering Generalities
- Transfer
- Testimonial
- Bandwagon

Societal Trust is a Network

- ◆ A Simplified Description of Societal Trust:
 - ◆ Weighted Directed Graph of Nodes and Weighted Arcs
 - Nodes = Societal Entities (People, Ideas, ...)
 - Arcs = Recommendation from an entity to another
 - Arc weight = Degree of entrustment
 - ◆ Then what is Propaganda?
 - Attempt to **modify the Trust Social Network** in ways beneficial to propagandist
 - ◆ And what is Web Spam?
 - Attempt to **modify the Web Graph** in ways beneficial to spammer



Web Spam as Propaganda

| SE's | Ranking | Spamming | Propaganda |
|---------|-------------------|-------------------------------|-------------------------|
| 1st Gen | Doc Similarity | Keyword stuffing | Glittering generalities |
| 2nd Gen | + Site popularity | + link farms | + Bandwagon |
| 3rd Gen | + Page reputation | + mutual admiration societies | + Testimonials |

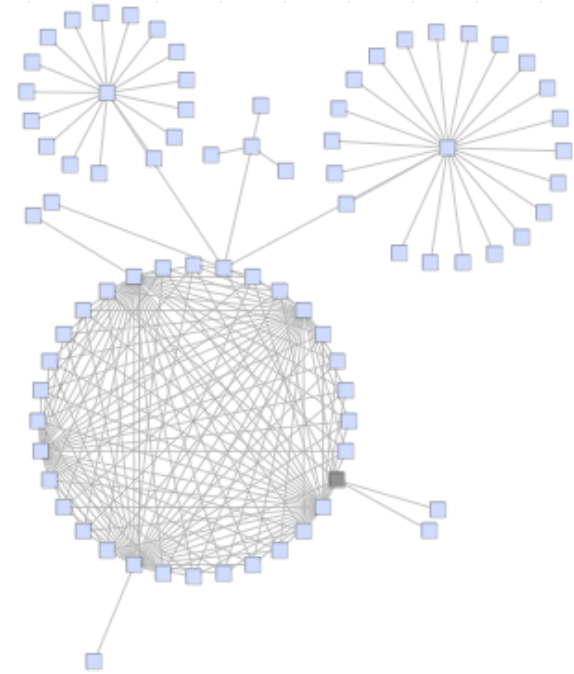
Web Spam is a major force behind Search Engine evolution

So what?

Can this understanding help us defend against web spam?

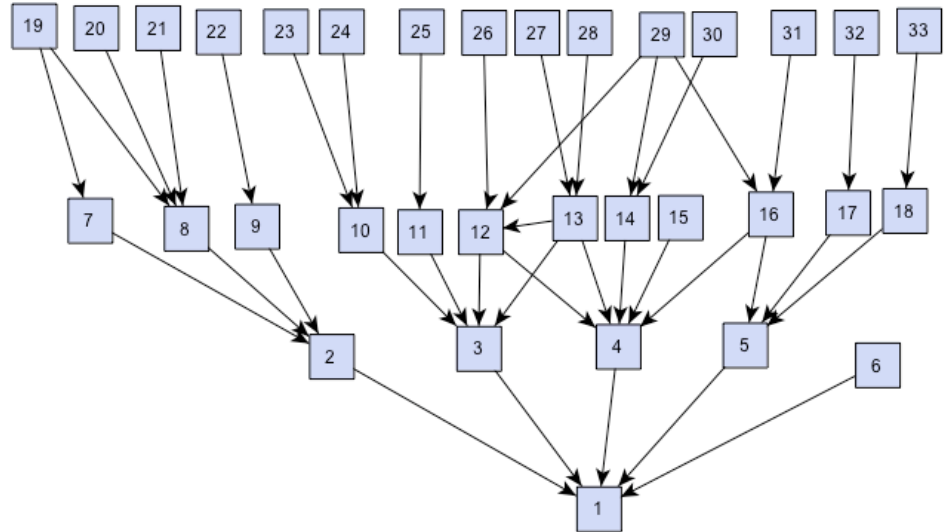
Anti-Propagandistic Lessons for Web

- ◆ How do you deal with propaganda in real life?
- ◆ **Backward propagation of distrust**
- ◆ *The recommender of an untrustworthy message becomes untrustworthy*
- ◆ Can you transfer this technique to the web?



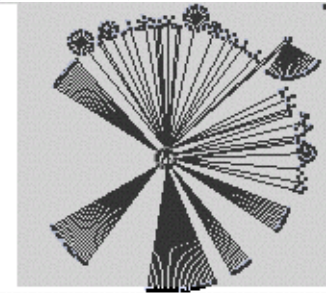
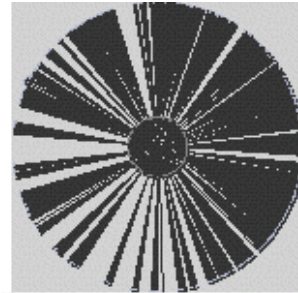
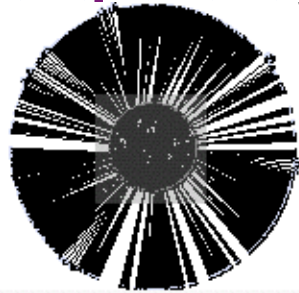
An Anti-Propagandistic Algorithm

- ◆ Start from untrustworthy site **s**
- ◆ $S = \{s\}$
- ◆ Using BFS for depth D do:
 - Find the set U of sites linking to sites in S
(using the Google API for up to B b-links/site)
 - Ignore blogs, directories, edu's
 - $S = S + U$
- ◆ Find the bi-connected component **BCC** of U that includes s



**BCC shows multiple paths
to boost the reputation of s**

Explored neighborhoods



- 0.3_http://www.ginkgo-bruba-online.com/
- 0.3_http://www.natural-enhancement-enlarge
- 0.3_http://www.nonijuce.us/
- 0.412_http://www.supergreen.biz/
- 0.416_http://www.health1nutrition.com/
- 0.422_http://www.amah.co.uk/
- 0.423_http://www.healthtreasures.net/
- 0.427_http://www.calcium-factor.com/
- 0.437_http://www.1for1.com/
- 0.43_http://www.vitaminwomen.com/
- 0.443_http://www.clear-lens-and-colored-co
- 0.445_http://www.phendimetrazine-director
- 0.446_http://www.vitabase.com/
- 0.451_http://www.biosync.us/
- 0.451_http://www.buy-viagra-free-prescripti
- 0.458_http://www.dressingyou.com/
- 0.459_http://www.cholesterol-metabolizer.cor
- 0.4_http://www.cancer-success.com/
- 0.4_http://www.vitaminstohealth.com/
- 0.517_http://www.juicing.com/
- 0.558_http://www.superior-coral.com/
- 0.571_http://www.arthritis-msm-supplements
- 0.584_http://www.coralcalciumetc.com/
- 0.604_http://www.bettercoralcalcium.com/
- 0.614_http://www.healthtreasures.com/
- 1.0_http://www.coral-calcium-benefits.com/

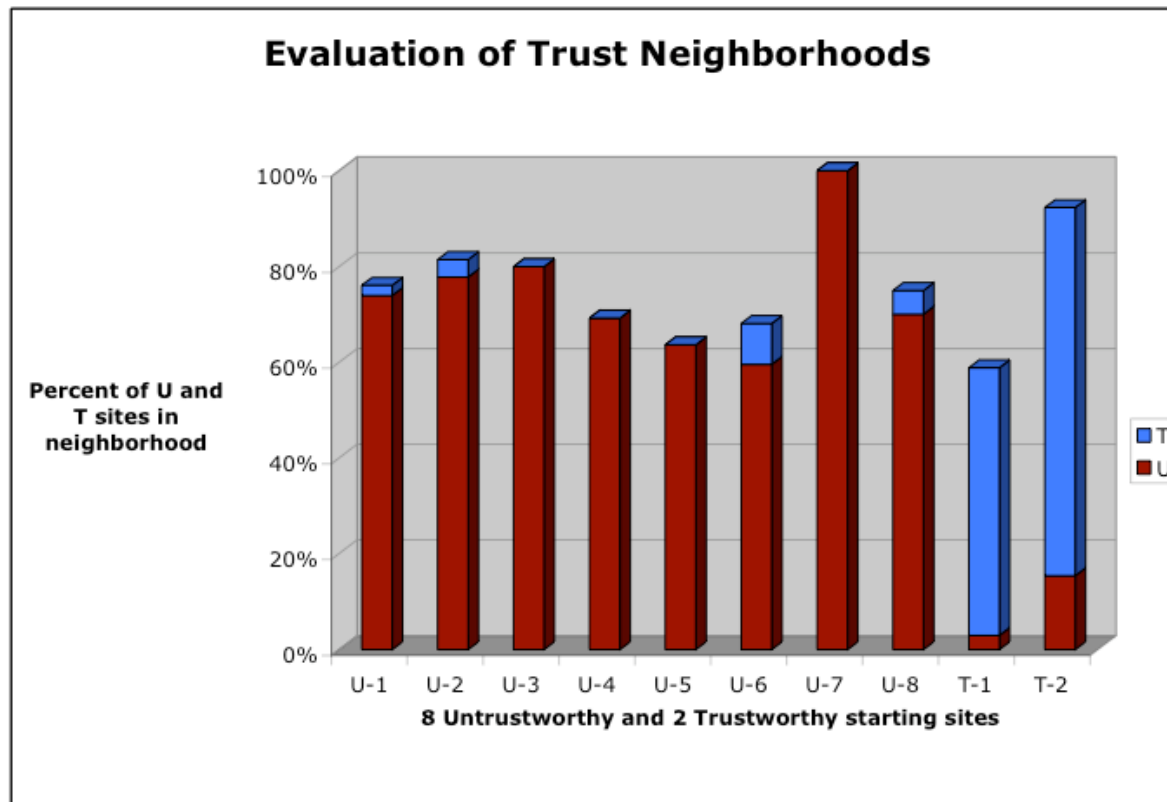
- 0.418_http://www.nutritionstreet.com/
- 0.42_http://www.newworldproducts.org/
- 0.42_http://www.smartwomensupplements.cor
- 0.433_http://www.supreme-greens-msm.org/
- 0.438_http://heartspring.net/
- 0.43_http://www.onlinecoralcalcium.com/
- 0.441_http://www.hgh-best-results.com/
- 0.443_http://www.healthadvancements.7p.com
- 0.44_http://www.internetarthritiscenter.com/
- 0.454_http://www.innerlifewellness.com/
- 0.455_http://www.utropin.com/
- 0.463_http://www.synflexonline.com/
- 0.466_http://www.cyber-supplements.com/
- 0.475_http://www.ultimate-orgasms-and-enha
- 0.486_http://www.greatestherbsonearth.com/
- 0.493_http://www.mens-health-naturally.com/
- 0.495_http://www.calcompnutrition.com/
- 0.501_http://www.hgh.nutritional-dietary-body
- 0.522_http://www.supergreen.biz/
- 0.54_http://www.health-information.biz/
- 0.553_http://www.skin-care-solutions.net/
- 0.58_http://healthproducts-usa.com/
- 0.5_http://vitaminmen.com/
- 0.5_http://www.amah.co.uk/
- 1.0_http://www.renuva.net/clinical.htm

- 0.300_http://www.mncss-universi
- 0.314_http://www.webmasterstop
- 0.316_http://www.weightlossguida
- 0.322_http://www.33homebizpar:
- 0.325_http://www.online-cialis.net
- 0.332_http://www.123-search-en
- 0.333_http://www.cyberflexing.ho
- 0.338_http://www.muscle-building
- 0.338_http://www.savoncoins.com
- 0.344_http://www.1-weight-loss-
- 0.353_http://www.fitnesszone.co.:
- 0.353_http://www.sarinoni.com/
- 0.358_http://www.diet-websites.c
- 0.362_http://www.voy.com/
- 0.364_http://www.1stopvitamin.cc
- 0.365_http://globalifestyle.net/
- 0.384_http://www.manandmollusc
- 0.387_http://www.trainer4you.con
- 0.396_http://weightlossandfitnessi
- 0.39_http://specialneedsfamilyfun
- 0.414_http://www.fitstep.com/
- 0.414_http://www.metaxa.net/
- 0.419_http://www.4hgh.ws/
- 0.429_http://burnthefat.2market4
- 0.453_http://www.coral1.com/
- 1.0_http://www.advice-hqh.com/

Evaluated Experimental Results

| Target | G | BCC | Trustworth | Untrstwrth | Directory |
|----------------------------|------|-----|-------------|--------------|-----------|
| renuva.net | 1307 | 228 | 2% = 1/46 | 74% = 34/46 | 13% |
| coral-calcium-benefits.com | 1380 | 266 | 4% = 2/54 | 78% = 42/54 | 7% |
| vespro.com | 875 | 97 | 0% = 0/20 | 80% = 16/20 | 15% |
| hardcorebodybuilding.com | 457 | 63 | 0% = 0/13 | 69% = 9/13 | 15% |
| maxsportsmag.com | 716 | 105 | 0% = 0/22 | 64% = 14/22 | 27% |
| coral1.com | 312 | 228 | 9% = 4/47 | 60% = 28/47 | 13% |
| genf20.com | 81 | 32 | 0% = 0/32 | 100% = 32/32 | 0% |
| 1stHGH.com | 1547 | 200 | 5% = 2/40 | 70% = 28/40 | 10% |
| hgfound.org | 1429 | 164 | 56% = 19/34 | 14% = 1/34 | 26% |
| advice-hgh.com | 241 | 13 | 77% = 10/13 | 15% = 2/13 | 8% |

Evaluated Experimental Results



Conclusions and Next Steps

- ◆ Web Spam / Cyberworld = Propaganda / Society
- ◆ Particular spamming techniques can be uncovered - then what?
- ◆ Spam becomes a necessity as web grows
 - “I spent all my life searching for the meaning of life...”
 - “If you cannot find it on eBay or Google, it does not exist”
- ◆ Spam to you, treasure to me
- ◆ *Who do you trust* is the right question to ask
and provide tools for managing trusted and distrusted
- ◆ Personalization of search
 - a search engine (component) per browser
 - Or: specialized search engines
- ◆ Education, critical thinking
 - What we believe, why we believe it
- ◆ Cyber-social structures and networks
 - I inherit the trusted/distrusted networks of the societies I join

How (not) To Solve The Problem

