SOCIAL SPAM DETECTION

Benjamin Markines
Ciro Cattuto
Filippo Menczer

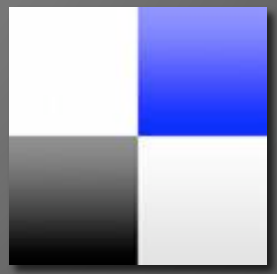
INDIANA UNIVERSITY
SCHOOL OF INFORMATICS
Bloomington

ISI Foundation

Social Applications

BibSonomy

myspace.com
a place for friends.

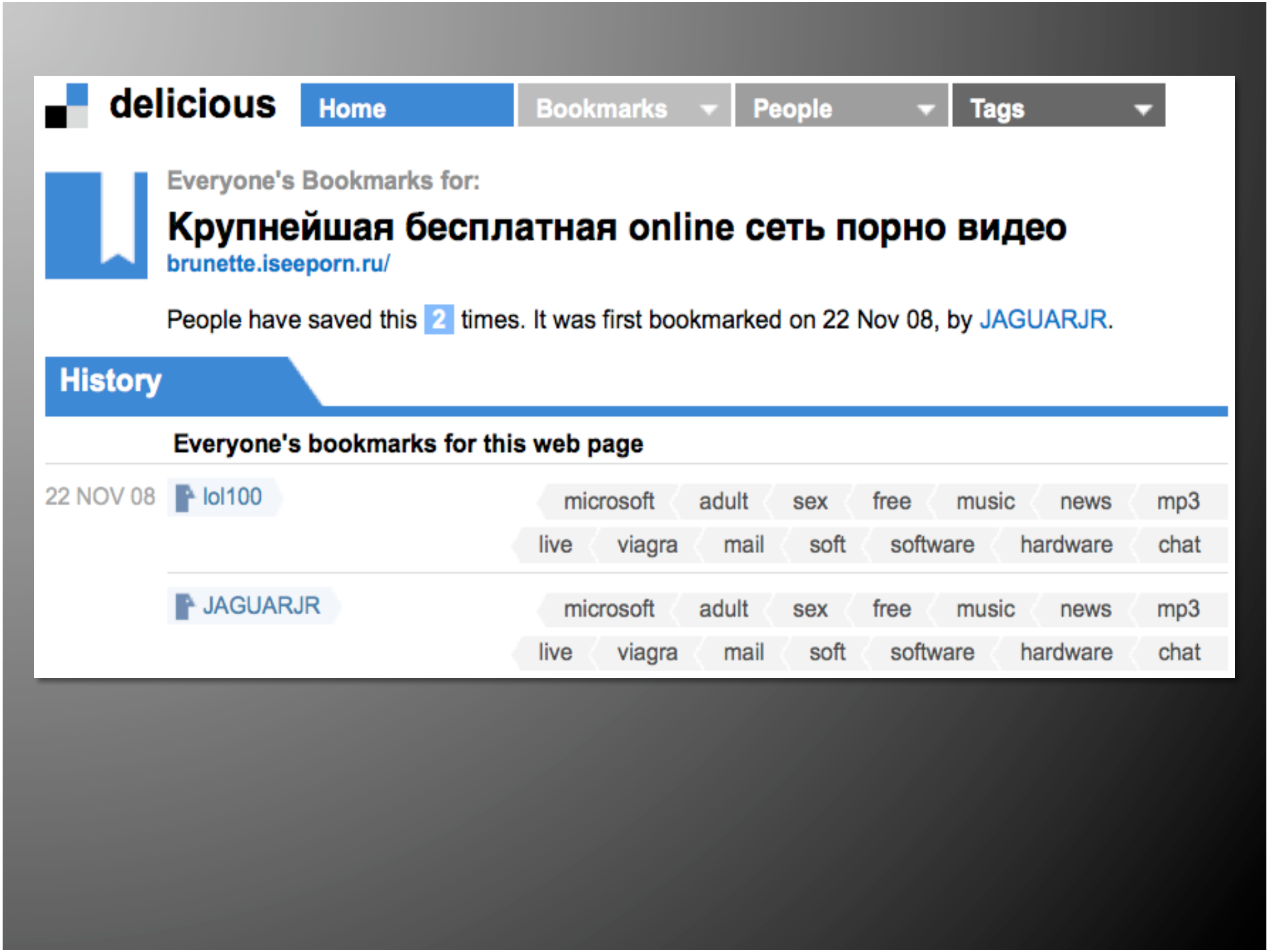


YouTube
Broadcast Yourself

LinkedIn

facebook

flickr


delicious
Home
Bookmarks
People
Tags
Everyone's Bookmarks for:
Крупнейшая бесплатная online сеть порно видео
brunette.iseeporn.ru/
People have saved this 2 times. It was first bookmarked on 22 Nov 08, by JAGUARJR.
History
Everyone's bookmarks for this web page
22 NOV 08
lol100
microsoft adult sex free music news mp3
live viagra mail soft software hardware chat
JAGUARJR
microsoft adult sex free music news mp3
live viagra mail soft software hardware chat

givealink.org

I donated my bookmarks to science. Did you?

Email: Register

Password: Log In

About: Project Help FAQ Privacy | **Links:** Home Share Manage Download RSS

Search

Recommendations

Enter URL

URL Search

Enter URL

Novel URL Search

Popular Tags

apartments buy cheap echo-ftp2 estate home house imported in india new office property real rent residential sale search shop test

Getting Started

personalized recommendations

social search

Top Ten Bookmarks ?

1. Заратустра. О бледном преступнике. | Мыслей нет, а книги есть.
2. Заратустра. О дереве на горе. | Мыслей нет, а книги есть.
3. Заратустра. О чтении и письме. | Мыслей нет, а книги есть.
4. Заратустра. О проповедниках смерти. | Мыслей нет, а книги есть.
5. Заратустра. О радостях и страстях. | Мыслей нет, а книги есть.
6. Фридрих Ницше. Речи заратустры. О трех превращениях. | Мыслей нет, а книги есть.
7. Предисловие заратустры. Ницше | Мыслей нет, а книги есть.
8. Принцеса на горошине и Бахчисарайский фонтан | Мыслей нет
9. Праотцы туалетов, особенно улыбнуло!

CandidOpinions.com

Most Candid Views, Opinions, Comments and Reviews about everything

« Desperate Housewives Coverage at Kikizo First videogames news,
Desperate Housewives Episode Guide – Desperate Housewives Season ... »

USATODAY.com - 'Housewives' sets up shop in computer

$99 - Custom Logo Design
Unique Designs, Unlimited Revisions Satisfaction or Moneyback Guarantee
LogoDesignTeam.com

Aaa Tree And Stump .com
10 % Off Competitor's Estimate ! Licensed - Insured - Worker's Comp.
aaatreeandstump.com

Make $30,100.71 In 2 Wks
Free report reveals hidden secrets to making a huge fortune online.
embassyinc.co.uk

Your Ad Here!
BIDVERTISER ADS

USATODAY.com – 'Housewives' sets up shop in computer game ... will announce a computer game based on ABC's mega-popular Desperate Housewives. ... And it wouldn't be a Desperate Housewives game if you couldn't try to steal a ...

YouTube – Desperate Housewives – Opening Credits Season 1
www.desperate-housewives.vip7.com ... desperate housewives wisteria lane teri hatcher marcia cross eva longoria (more) (less) ...

This entry was posted on Saturday, September 22nd, 2007 at 6:05 pm and is filed under TV. You can follow any responses to this entry through the RSS 2.0 feed. You can leave a response, or trackback from your own site.

Leave a Reply

Name (required)

Mail (will not be published) (required)

Aaa Tree And Stump .com
10 % Off Competitor's Estimate ! Licensed - Insured - Worker's Comp.
aaatreeandstump.com

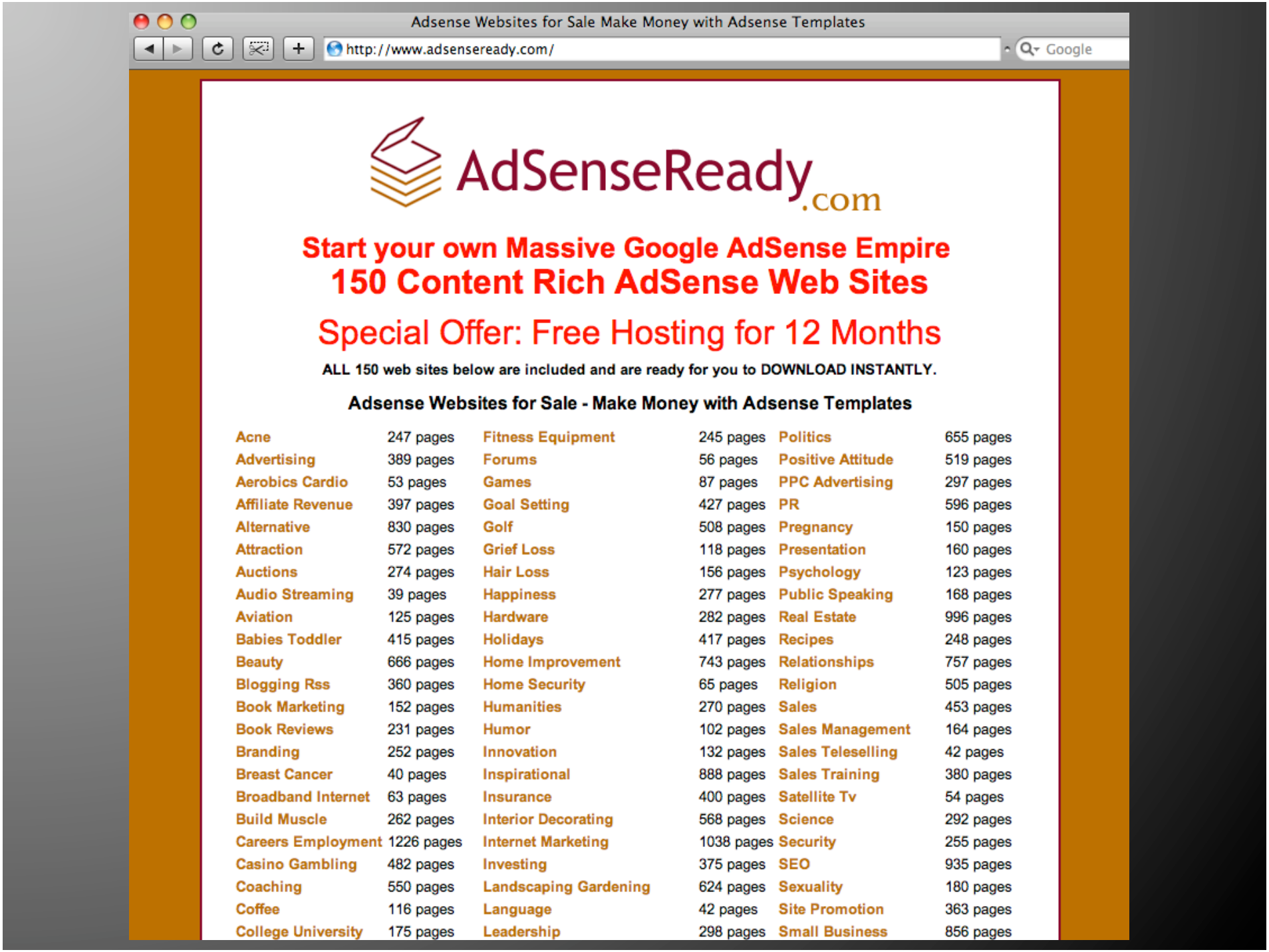
Make $30,100.71 In 2 Wks
Free report reveals hidden secrets to making a huge fortune online.
embassyinc.co.uk

Free Credit Repair Info
Everything you need to know about your credit is yours for the taking
creditinfocenter.com

Last Chaos Online Game
Free to Play, Pets, Guilds and War, Great Graphics-role-playing quests.
lastchaos-usa.com

Make $100 - $3.500 Daily
Let the system do 99% of the work. Learn the potential of Cash Gifting.
epicwealthsystems.com

Your Ad Here!
BIDVERTISER ADS


Adsense Websites for Sale Make Money with Adsense Templates
http://www.adsenseready.com/
Google
AdSenseReady.com
Start your own Massive Google AdSense Empire
150 Content Rich AdSense Web Sites
Special Offer: Free Hosting for 12 Months
ALL 150 web sites below are included and are ready for you to DOWNLOAD INSTANTLY.
Adsense Websites for Sale - Make Money with Adsense Templates
Acne 247 pages
Advertising 389 pages
Aerobics Cardio 53 pages
Affiliate Revenue 397 pages
Alternative 830 pages
Attraction 572 pages
Auctions 274 pages
Audio Streaming 39 pages
Aviation 125 pages
Babies Toddler 415 pages
Beauty 666 pages
Blogging Rss 360 pages
Book Marketing 152 pages
Book Reviews 231 pages
Branding 252 pages
Breast Cancer 40 pages
Broadband Internet 63 pages
Build Muscle 262 pages
Careers Employment 1226 pages
Casino Gambling 482 pages
Coaching 550 pages
Coffee 116 pages
College University 175 pages
Fitness Equipment 245 pages
Forums 56 pages
Games 87 pages
Goal Setting 427 pages
Golf 508 pages
Grief Loss 118 pages
Hair Loss 156 pages
Happiness 277 pages
Hardware 282 pages
Holidays 417 pages
Home Improvement 743 pages
Home Security 65 pages
Humanities 270 pages
Humor 102 pages
Innovation 132 pages
Inspirational 888 pages
Insurance 400 pages
Interior Decorating 568 pages
Internet Marketing 1038 pages
Investing 375 pages
Landscaping Gardening 624 pages
Language 42 pages
Leadership 298 pages
Politics 655 pages
Positive Attitude 519 pages
PPC Advertising 297 pages
PR 596 pages
Pregnancy 150 pages
Presentation 160 pages
Psychology 123 pages
Public Speaking 168 pages
Real Estate 996 pages
Recipes 248 pages
Relationships 757 pages
Religion 505 pages
Sales 453 pages
Sales Management 164 pages
Sales Teleselling 42 pages
Sales Training 380 pages
Satellite Tv 54 pages
Science 292 pages
Security 255 pages
SEO 935 pages
Sexuality 180 pages
Site Promotion 363 pages
Small Business 856 pages

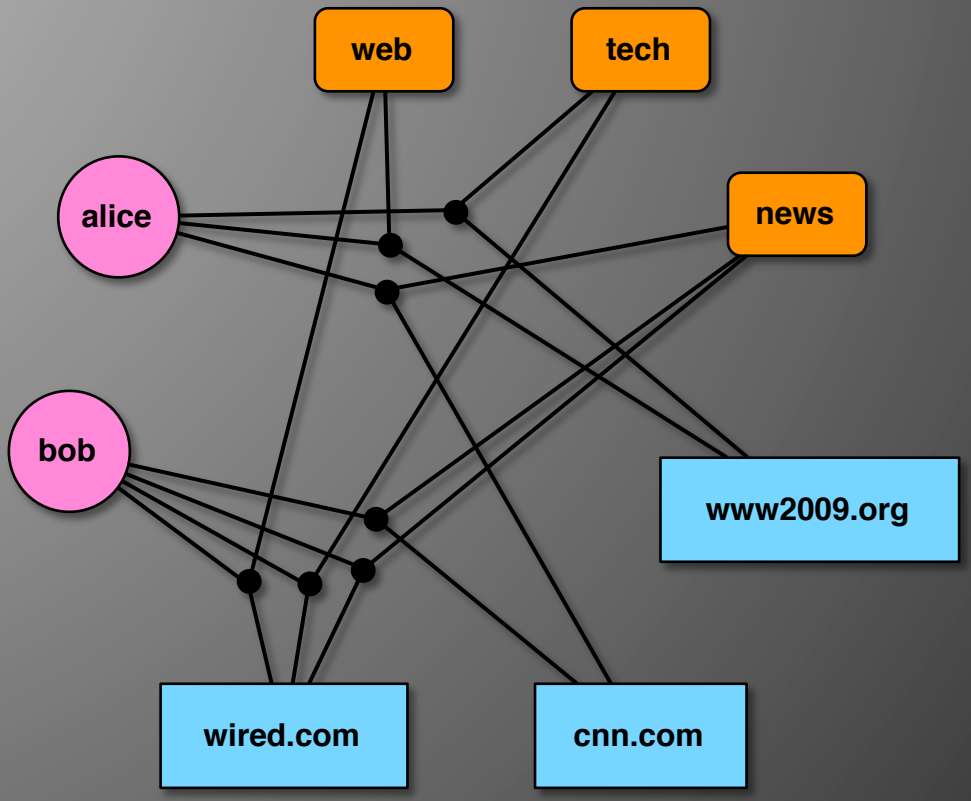
Who Wins? Who Loses?

- Beneficiaries
  - Spammer
  - Intermediary
- Non-beneficiaries
  - Systems: search engines, tagging
  - Information consumers
  - Original authors
- ?
  - Surfer
  - Advertisers

Agenda

- **Features**
  - **Folksonomy description**
  - **Feature descriptions**
    - **post level**
    - **resource level**
    - **user level**
- Feature analysis
  - Dataset description
- Social spam detection

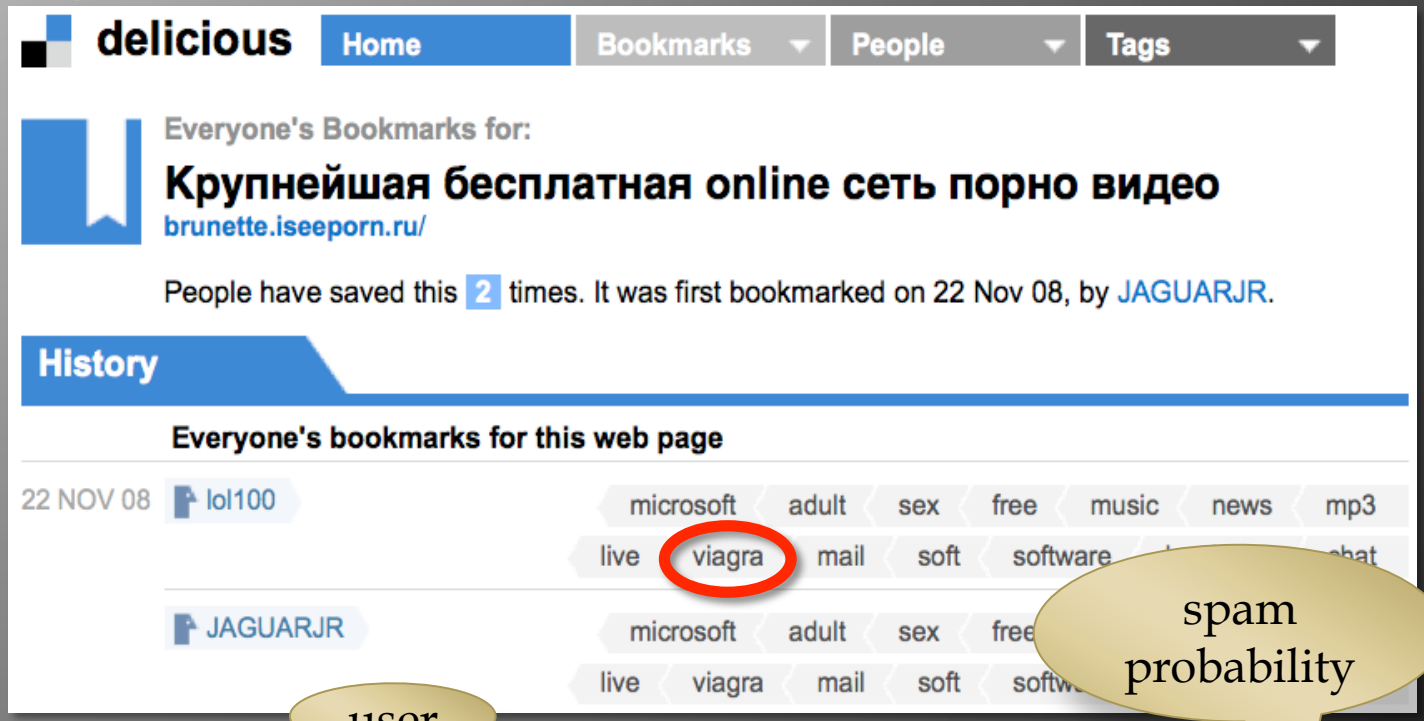
Folksonomy Definition



$$F = (U, T, R, Y), Y \subseteq U \times T \times R \text{ (the triples)}$$

Features: Post Level

- TagSpam



$$f_{TagSpam}(u,r) = \frac{1}{|T(u,r)|} \sum_{t \in T(u,r)} \Pr(t)$$

user

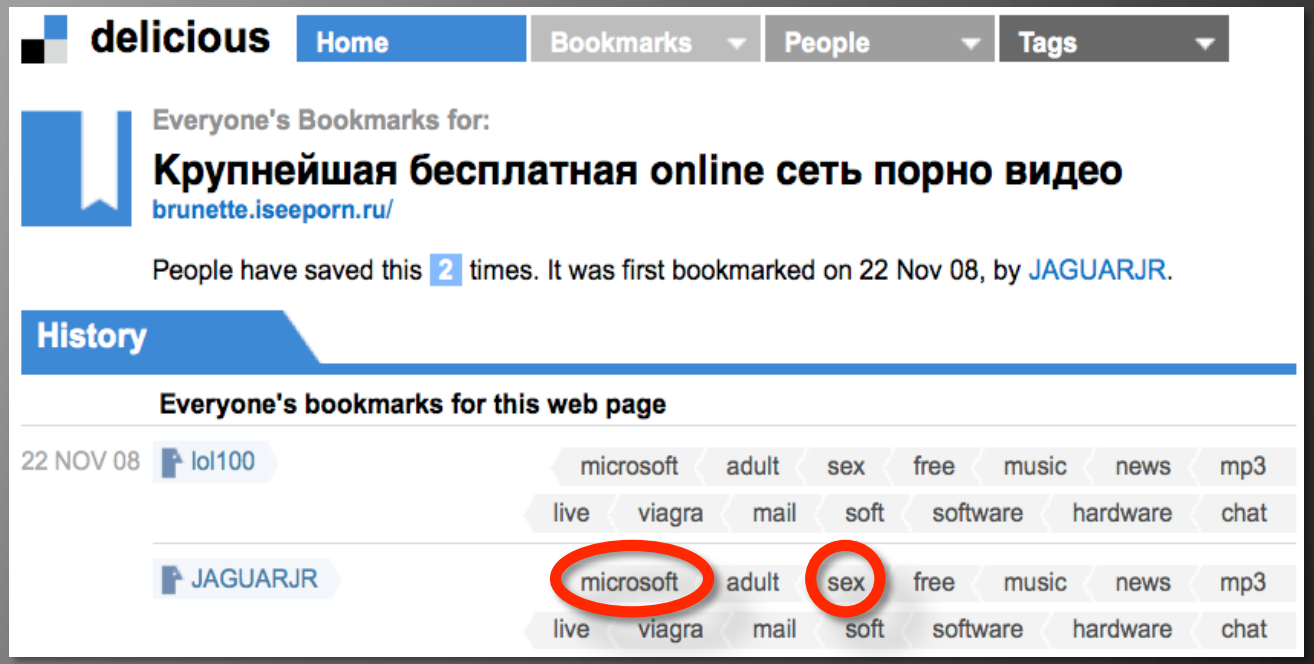
resource

post tags

spam probability

Features: Post Level

- TagBlur

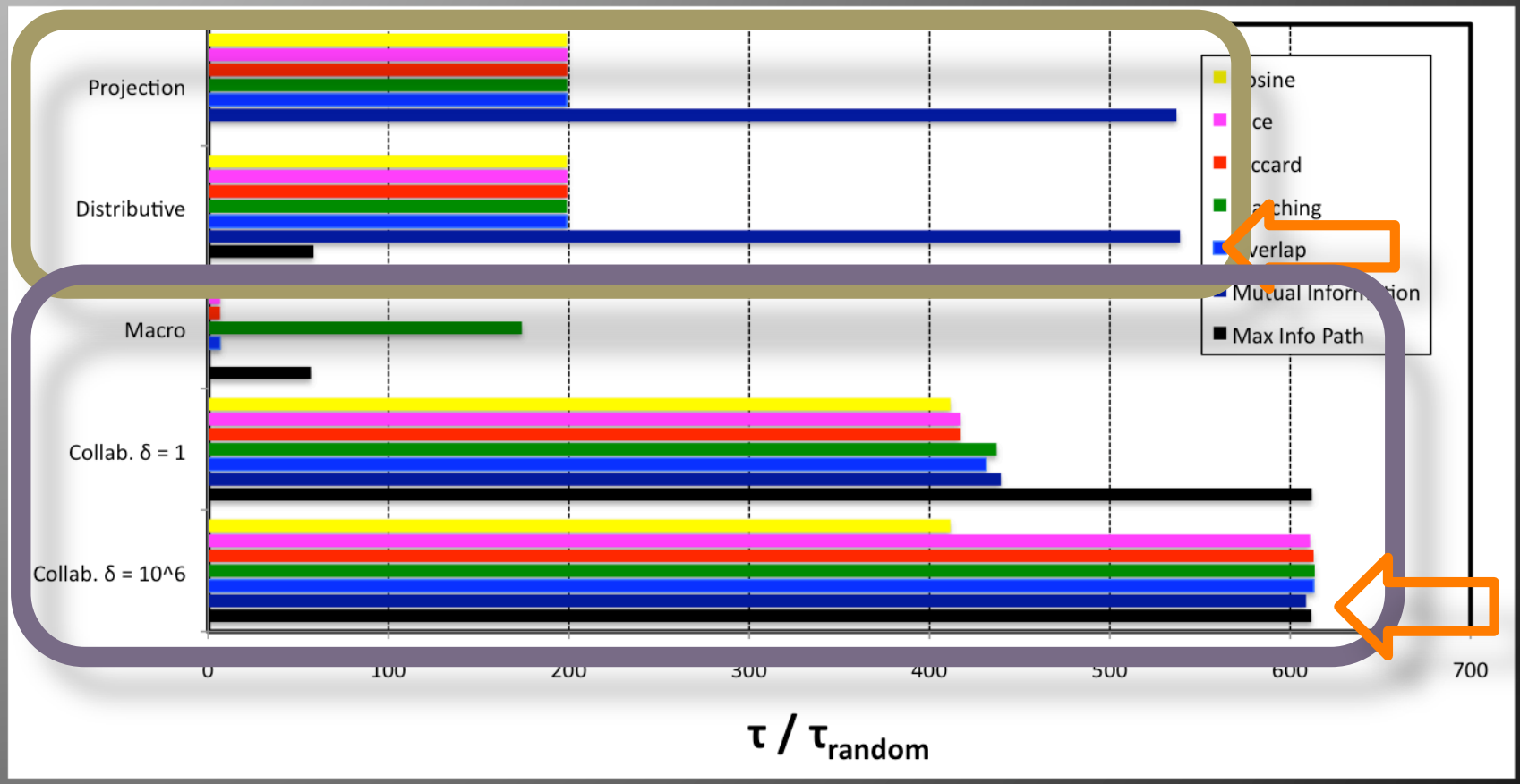


$$f_{TagBlur}(u,r) = \frac{1}{Z} \sum_{t_1 \neq t_2 \in T(u,r)} \frac{1}{\sigma(t_1,t_2)+\epsilon} - \frac{1}{1+\epsilon}$$

| tag pairs |

tag similarity

Tag Similarity



random $\tau = 10^{-4}$

WWW 2009
HT 2009

Features: Resource Level

- DomFp

shingle similarity

DOM fingerprint

spam probability

$$f_{DomFp}(r) = \frac{\sum_{k \in K} \sigma(k(r), k) \cdot \Pr(k)}{\sum_{k \in K} \sigma(k(r), k)}$$

Features: Resource Level

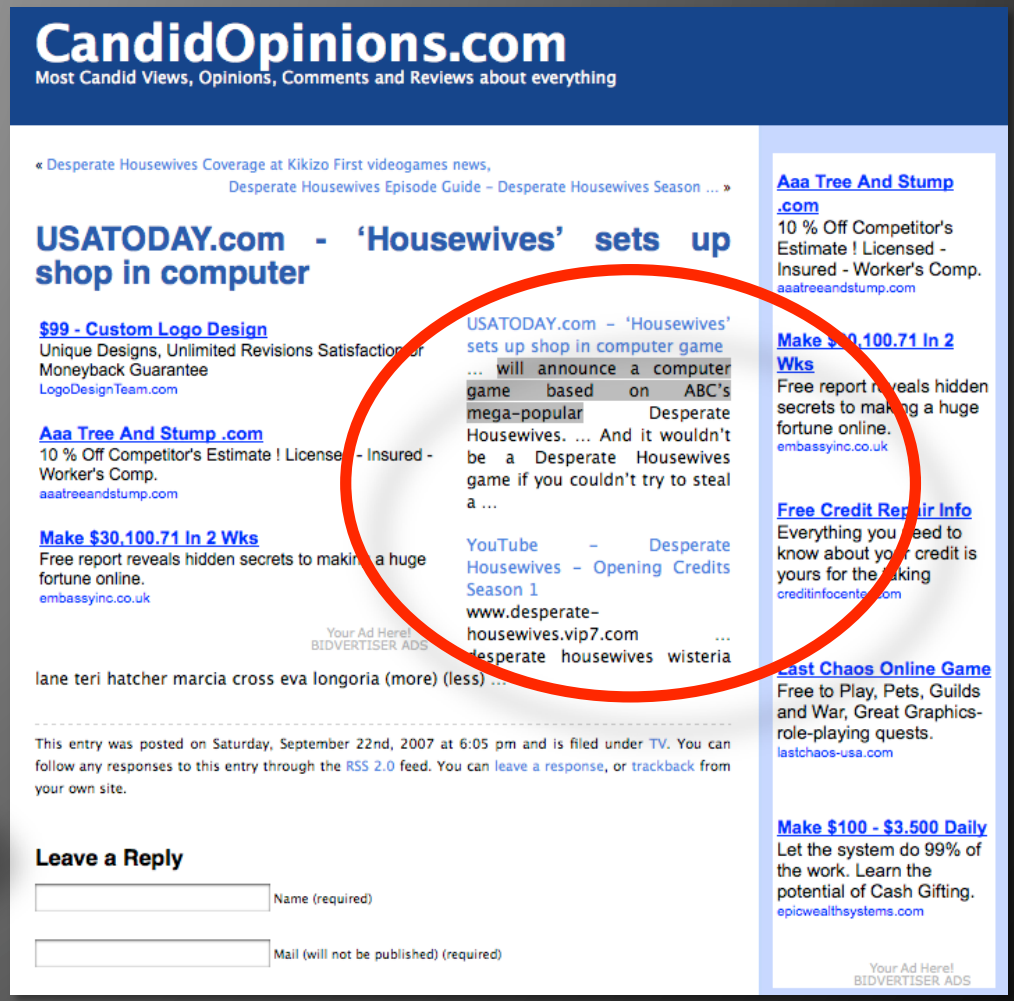
- NumAds

number of ads

$$f_{NumAds}(r) = g(r)/g_{max}$$

Features: Resource Level

- Plagiarism

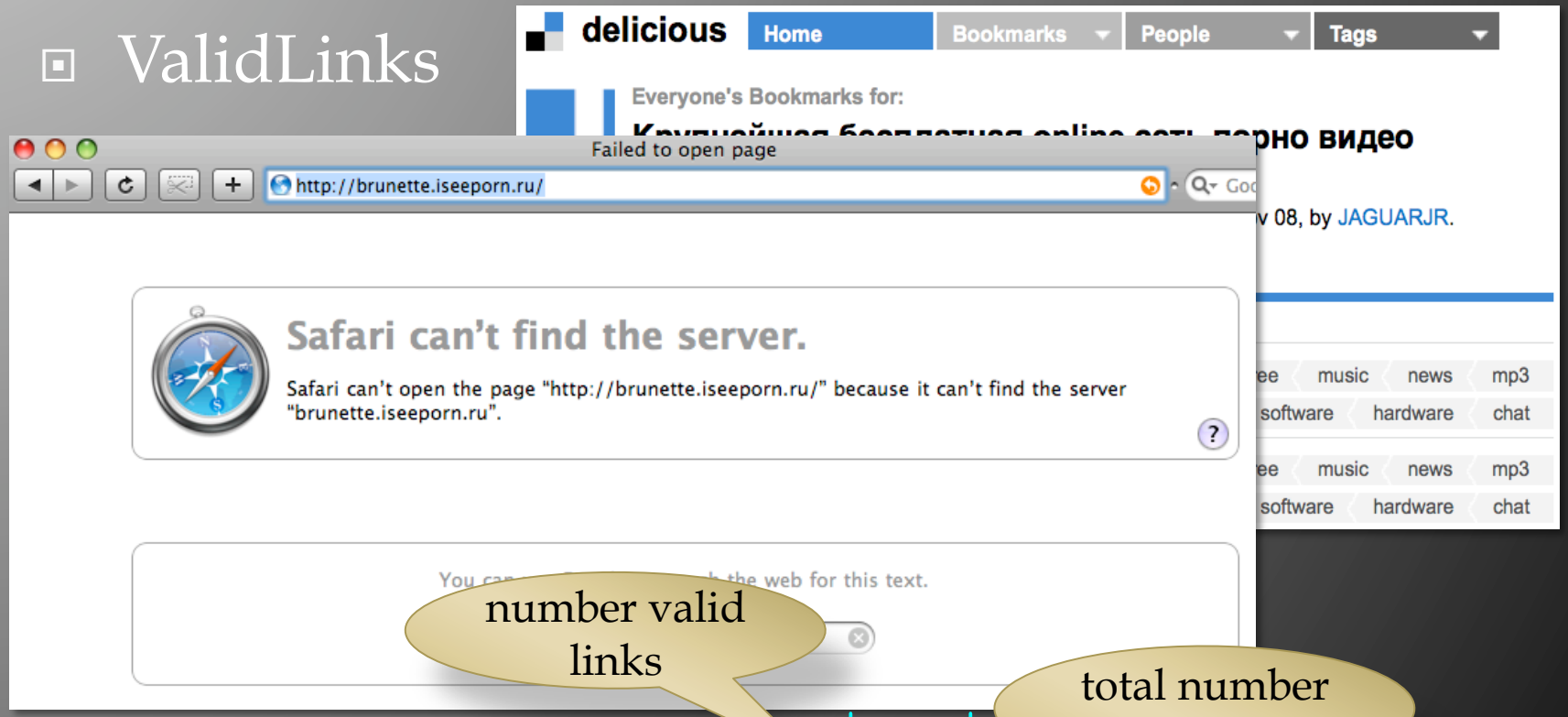


number of more authoritative sources matching exact phrase

$$f_{Plagiarism}(r) = y(r)/y_{max}$$

Features: User Level

- ValidLinks



number valid links

total number of links

$$f_{ValidLinks}(u) = \frac{|V_u|}{|R_u|}$$

BibSonomy Dataset

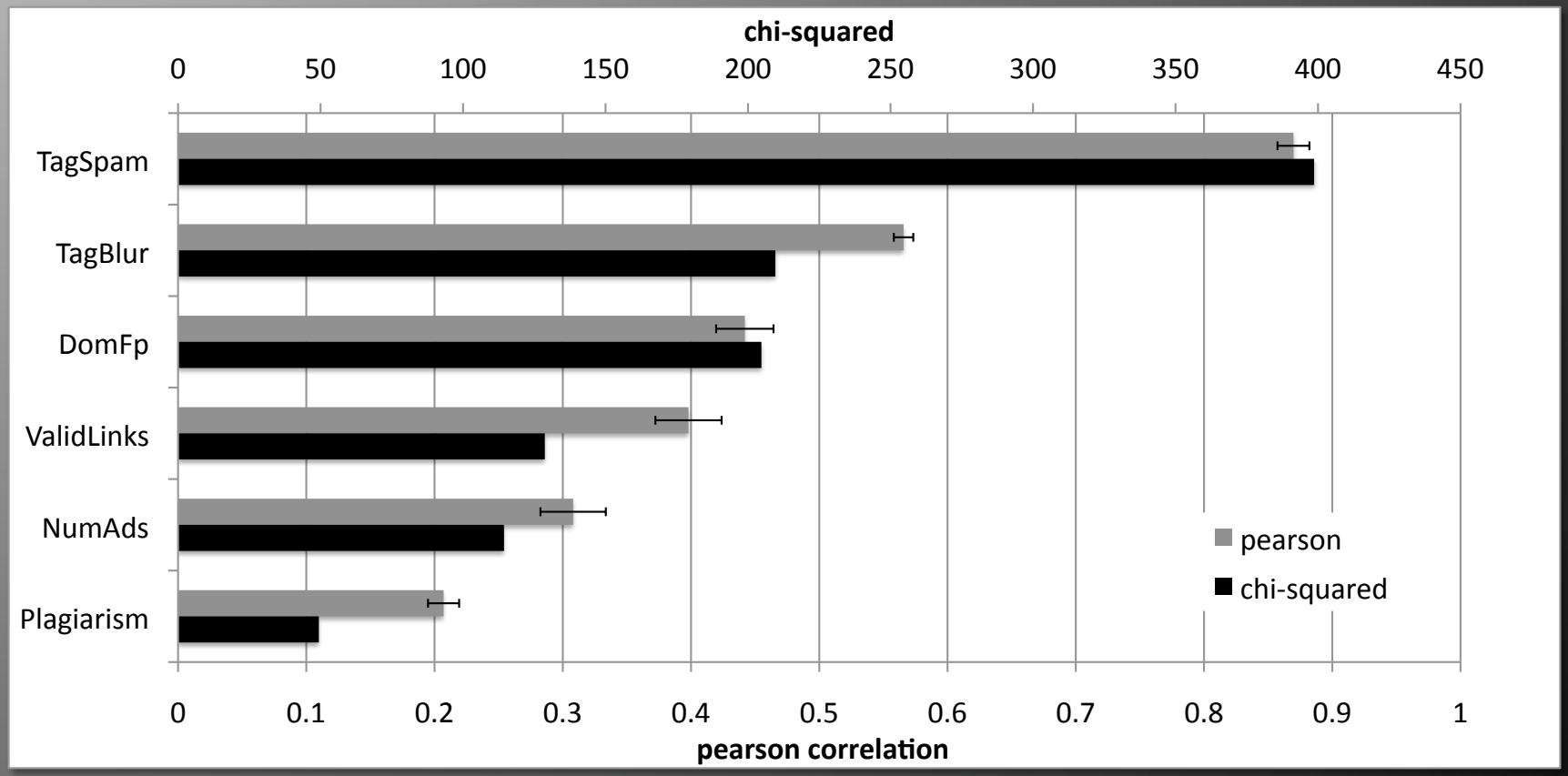
- BibSonomy.org
- Spam is labeled at user level
  - Aggregate for user level

$$f(u) = \frac{1}{|P(u)|} \sum_{(u,r) \in P(u)} f(u, r)$$

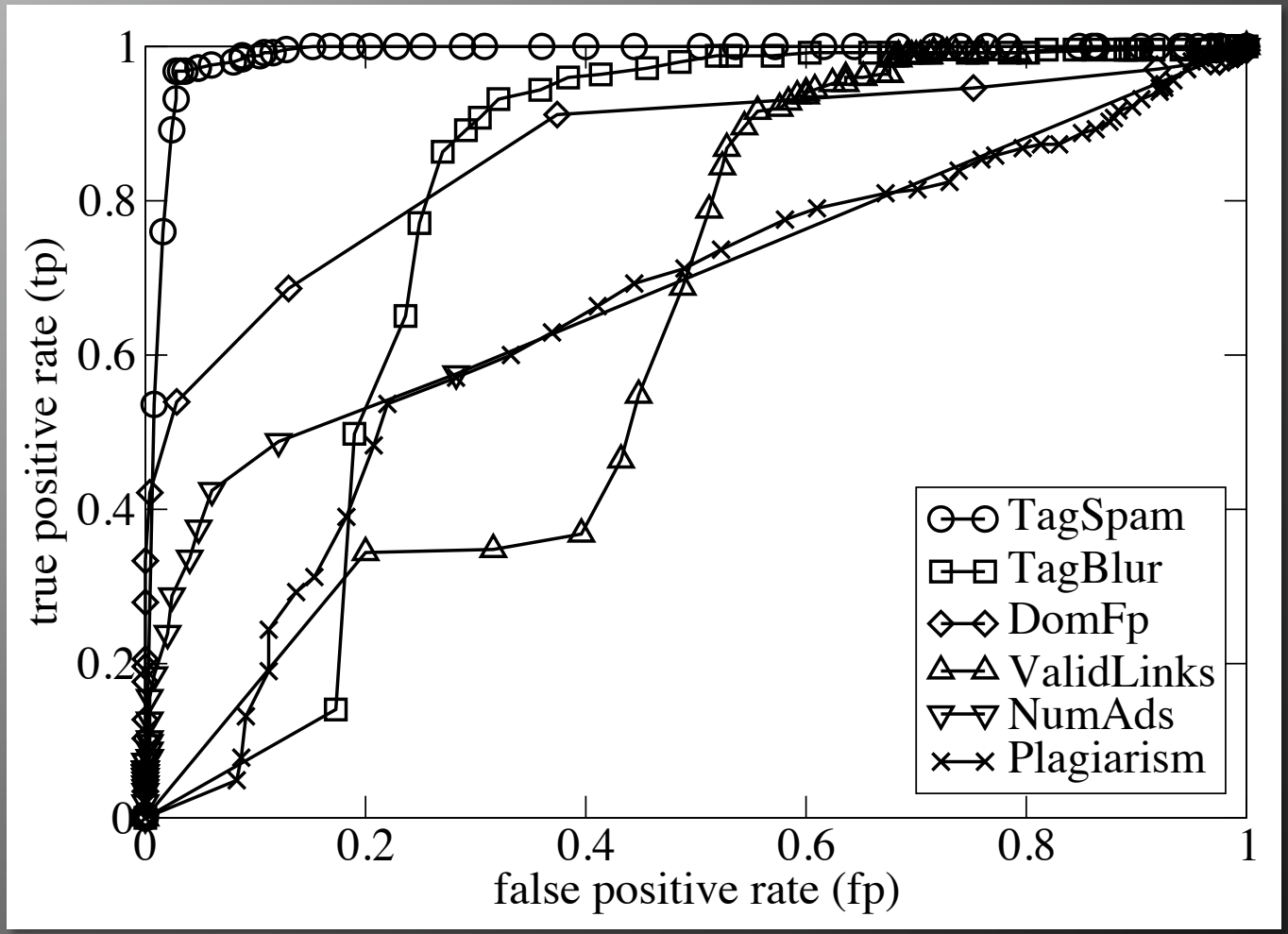
posts

- Sampled 1000 users
  - 500 spammers
- 500 users in training set/test set
  - 250 spammers

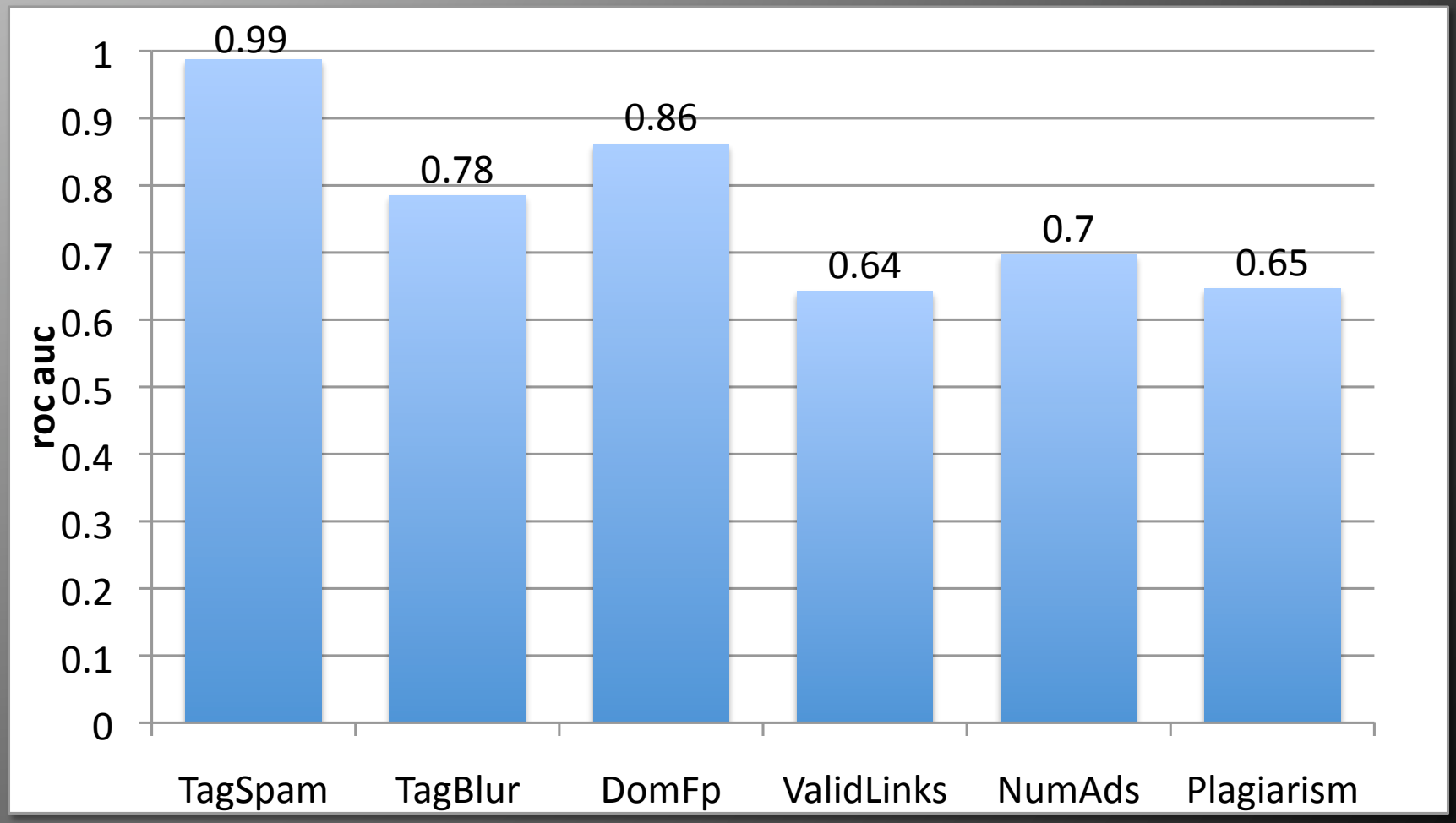
Feature Analysis: Discrimination Power



ROC Curves


true positive rate (tp)
false positive rate (fp)
0
0.2
0.4
0.6
0.8
1
TagSpam
TagBlur
DomFp
ValidLinks
NumAds
Plagiarism

AUC: Area Under ROC Curve



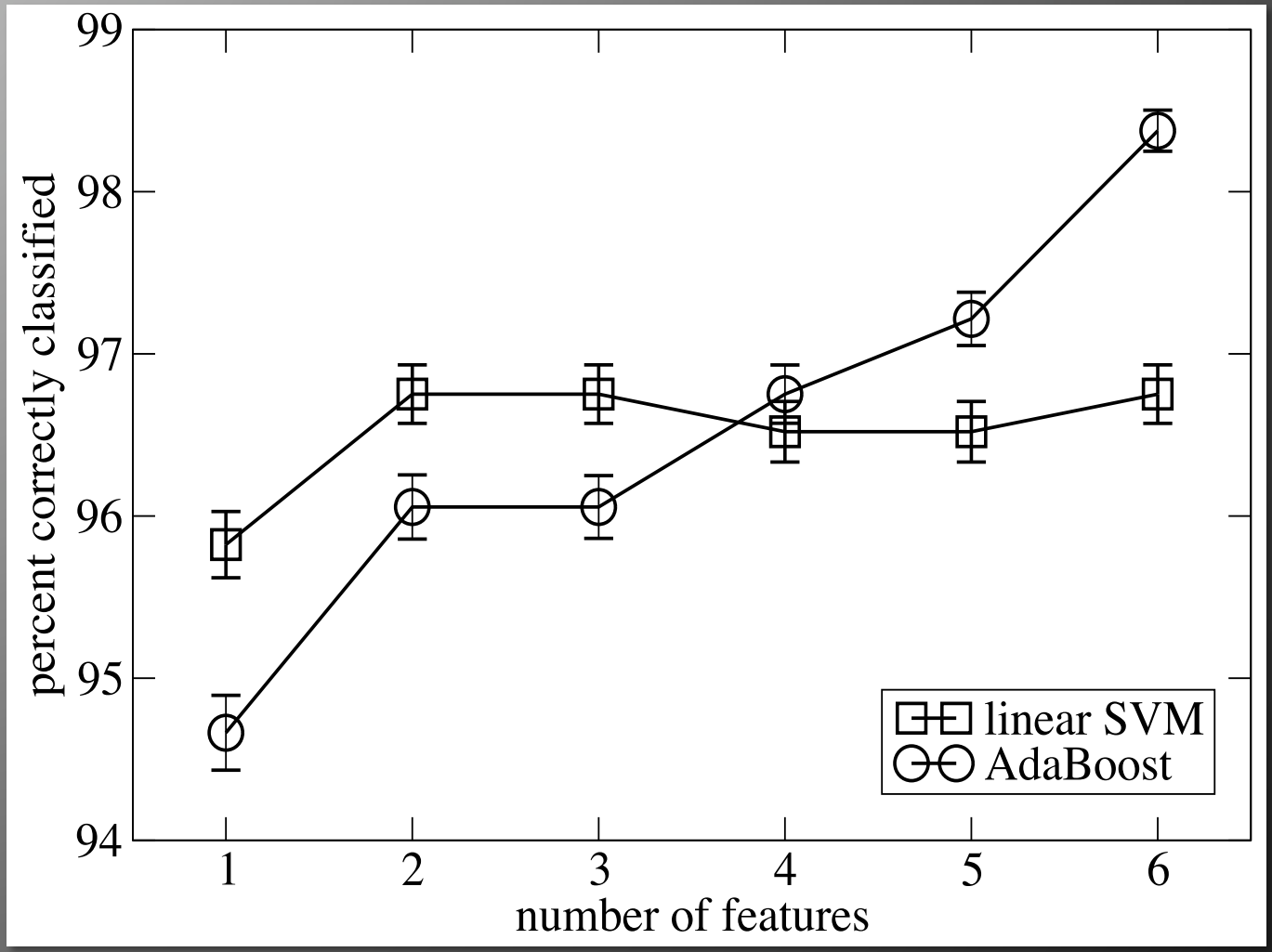
Social Spam Detection

Weka Classifier	Accuracy	FP	F_1
LogitBoost	**97.91%**	.018	**.978**
LWL	97.68%	**.013**	.975
AdaBoostM1	97.68%	.018	.975
ConjunctiveRule	97.68%	**.013**	.975
DecisionTable	97.68%	.018	.975
DecisionStump	97.68%	**.013**	.975
RandomCommittee	97.45%	.018	.973
RandomForest	97.45%	.018	.973
Bagging	97.22%	.022	.970
NNge	97.22%	.022	.970

Social Spam Detection

Features	linear SVM			AdaBoost		
	Accuracy	FP	F_1	Accuracy	FP	F_1
TagSpam	95.82%	.061	.957	94.66%	.048	.943
+ TagBlur	96.75%	.048	.966	96.06%	.044	.958
+ DomFp	96.75%	.048	.966	96.06%	.044	.958
+ ValidLinks	96.52%	.048	.964	96.75%	.026	.965
+ NumAds	96.52%	.048	.964	97.22%	.026	.970
+ Plagiarism	96.75%	.048	.966	**98.38%**	**.022**	**.983**

Social Spam Detection



Related Work

- Web/Email Spam
  - Attenberg and Suel 2008, Gyöngyi et al. 2004
- Social Spam
  - Heymann et al. 2007
- Spam Detection
  - Krause et al. 2008, Caverlee et al. 2008, Benevenuto et al. 2008, Koutrika et al. 2007
- ECML PKDD Discovery Challenge 2008
  - Held by BibSonomy team
  - Gkanogiannis and Kalamboukis 2008, Chevalier and Gramme 2008, Kim and Hwang 2008

Conclusions

- Identified/analyzed 6 features for spam detection
  - TagSpam alone achieves 0.99 ROC AUC outperforming ECML PKDD Discovery Challenge 2008
- Accuracy over 98% with AdaBoost
  - False-positive rate: 0.022
  - These results set the state of the art
    - could improve by combining with other features, e.g. Krause et al. 2008
- Limitations
  - Efficiency issues
  - Bootstrap issues

Feature Issues

- TagSpam
  - Depends on a set of labeled tags
- TagBlur
  - Depends on a notion of similarity/distances
  - Assumes spam does not dominate the folksonomy, affecting distances
- DomFp
  - Depends on a set of labeled fingerprints
  - Requires page download
- NumAds
  - Requires page download
- Plagiarism
  - Requires page download
  - Search engine cooperation
- ValidLinks
  - HEAD request per resource

THANK YOU!

- **Features**
  - **Folksonomy description**
  - **Feature descriptions**
    - **post level**
    - **resource level**
    - **user level**
- Feature analysis
- Social spam detection

Benjamin Markines

Ciro Cattuto

Filippo Menczer

BibSonomy Team

http://www.bibsonomy.org