SOCIAL SPAM DETECTION

Benjamin Markines
Ciro Cattuto
Filippo Menczer
Social Applications

BibSonomy

givealink.org
I donated my bookmarks to science.

myspace.com
a place for friends...

LinkedIn

Facebook

YouTube

Broadcast Yourself

flickr
Everyone's Bookmarks for:
Крупнейшая бесплатная online сеть порно видео
brunette.iseporn.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

<table>
<thead>
<tr>
<th>Date</th>
<th>Username</th>
<th>Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 Nov 08</td>
<td>lol100</td>
<td>microsoft, adult, sex, free,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>music, news, mp3, live, viagra, mail, soft, software, hardware, chat</td>
</tr>
<tr>
<td></td>
<td>JAGUARJR</td>
<td>microsoft, adult, sex, free,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>music, news, mp3, live, viagra, mail, soft, software, hardware, chat</td>
</tr>
</tbody>
</table>
Getting Started

donate your bookmarks

share the missing links

personalized recommendations

social search

Top Ten Bookmarks

1. Заратустра. О бедном 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 заутрусты. О трех превращениях. | Мыслей нет, а книги есть.
7. Предисловие заратустры. Ницше | Мыслей нет, а книги есть.
8. Принцесса на горошине и Бахчисарайский фонтан | Мыслей нет, а книги есть.
9. Праотцы турелотов, особенне улубнуло!
USATODAY.com - ‘Housewives’ sets up shop in computer

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.
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.

<table>
<thead>
<tr>
<th>Acne</th>
<th>247 pages</th>
<th>Fitness Equipment</th>
<th>245 pages</th>
<th>Politics</th>
<th>855 pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advertising</td>
<td>380 pages</td>
<td>Forums</td>
<td>56 pages</td>
<td>Positive Attitude</td>
<td>519 pages</td>
</tr>
<tr>
<td>Aerobics Cardio</td>
<td>53 pages</td>
<td>Games</td>
<td>87 pages</td>
<td>PPC Advertising</td>
<td>297 pages</td>
</tr>
<tr>
<td>Affiliate Revenue</td>
<td>337 pages</td>
<td>Goal Setting</td>
<td>427 pages</td>
<td>PR</td>
<td>596 pages</td>
</tr>
<tr>
<td>Alternativo</td>
<td>830 pages</td>
<td>Golf</td>
<td>508 pages</td>
<td>Pregnancy</td>
<td>150 pages</td>
</tr>
<tr>
<td>Attraction</td>
<td>572 pages</td>
<td>Grief Loss</td>
<td>118 pages</td>
<td>Presentation</td>
<td>160 pages</td>
</tr>
<tr>
<td>Auctions</td>
<td>274 pages</td>
<td>Hair Loss</td>
<td>156 pages</td>
<td>Psychology</td>
<td>120 pages</td>
</tr>
<tr>
<td>Audio Streaming</td>
<td>39 pages</td>
<td>Happiness</td>
<td>277 pages</td>
<td>Public Speaking</td>
<td>168 pages</td>
</tr>
<tr>
<td>Aviation</td>
<td>125 pages</td>
<td>Hardware</td>
<td>282 pages</td>
<td>Real Estate</td>
<td>996 pages</td>
</tr>
<tr>
<td>Babies Toddler</td>
<td>415 pages</td>
<td>Holidays</td>
<td>417 pages</td>
<td>Recipes</td>
<td>248 pages</td>
</tr>
<tr>
<td>Beauty</td>
<td>666 pages</td>
<td>Home Improvement</td>
<td>743 pages</td>
<td>Relationships</td>
<td>757 pages</td>
</tr>
<tr>
<td>Blogging Ras</td>
<td>350 pages</td>
<td>Home Security</td>
<td>65 pages</td>
<td>Religion</td>
<td>505 pages</td>
</tr>
<tr>
<td>Book Marketing</td>
<td>152 pages</td>
<td>Humanities</td>
<td>270 pages</td>
<td>Sales</td>
<td>453 pages</td>
</tr>
<tr>
<td>Book Reviews</td>
<td>231 pages</td>
<td>Humor</td>
<td>102 pages</td>
<td>Sales Management</td>
<td>184 pages</td>
</tr>
<tr>
<td>Branding</td>
<td>252 pages</td>
<td>Innovation</td>
<td>132 pages</td>
<td>Sales Teleselling</td>
<td>42 pages</td>
</tr>
<tr>
<td>Breast Cancer</td>
<td>40 pages</td>
<td>Inspirational</td>
<td>888 pages</td>
<td>Sales Training</td>
<td>380 pages</td>
</tr>
<tr>
<td>Broadband Internet</td>
<td>63 pages</td>
<td>Insurance</td>
<td>400 pages</td>
<td>Satellite Tv</td>
<td>54 pages</td>
</tr>
<tr>
<td>Build Muscle</td>
<td>262 pages</td>
<td>Interior Decorating</td>
<td>568 pages</td>
<td>Science</td>
<td>292 pages</td>
</tr>
<tr>
<td>Careers Employment</td>
<td>1226 pages</td>
<td>Internet Marketing</td>
<td>1036 pages</td>
<td>Security</td>
<td>255 pages</td>
</tr>
<tr>
<td>Casino Gambling</td>
<td>482 pages</td>
<td>Investing</td>
<td>375 pages</td>
<td>SEO</td>
<td>935 pages</td>
</tr>
<tr>
<td>Coaching</td>
<td>550 pages</td>
<td>Landscaping Gardening</td>
<td>624 pages</td>
<td>Sexuality</td>
<td>180 pages</td>
</tr>
<tr>
<td>Coffee</td>
<td>116 pages</td>
<td>Language</td>
<td>42 pages</td>
<td>Site Promotion</td>
<td>363 pages</td>
</tr>
<tr>
<td>College University</td>
<td>175 pages</td>
<td>Leadership</td>
<td>298 pages</td>
<td>Small Business</td>
<td>856 pages</td>
</tr>
</tbody>
</table>
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
$F = (U, T, R, Y), \ Y \subseteq U \times T \times R \ (\text{the triples})$
Features: Post Level

- TagSpam

\[ f_{\text{TagSpam}}(u, r) = \frac{1}{|T(u, r)|} \sum_{t \in T(u, r)} \Pr(t) \]
Features: Post Level

- TagBlur

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

\[|\text{tag pairs}| \]

\[\text{tag similarity}\]
Tag Similarity

random $\tau = 10^{-4}$
Features: Resource Level

- DomFp

\[ f_{DomFp}(r) = \frac{\sum_{k \in K} \sigma(k(r), k) \cdot \Pr(k)}{\sum_{k \in K} \sigma(k(r), k)} \]
Features: Resource Level

- NumAds

\[ f_{\text{NumAds}}(r) = \frac{g(r)}{g_{\text{max}}} \]
Features: Resource Level

- Plagiarism

\[ f_{\text{Plagiarism}}(r) = \frac{y(r)}{y_{\text{max}}} \]
ValidLinks

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

- number valid links
- total number of links
BibSonomy Dataset

- BibSonomy.org
- Spam is labeled at user level
  - Aggregate for user level
    \[ f(u) = \frac{1}{|P(u)|} \sum_{(u,r) \in P(u)} f(u, r) \]
- Sampled 1000 users
  - 500 spammers
- 500 users in training set/test set
  - 250 spammers
Feature Analysis: Discrimination Power

- TagSpam
- TagBlur
- DomFp
- ValidLinks
- NumAds
- Plagiarism

chi-squared
pearson correlation
ROC Curves

true positive rate (tp)

false positive rate (fp)

- TagSpam
- TagBlur
- DomFp
- ValidLinks
- NumAds
- Plagiarism
AUC: Area Under ROC Curve

<table>
<thead>
<tr>
<th>Feature</th>
<th>ROC AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>TagSpam</td>
<td>0.99</td>
</tr>
<tr>
<td>TagBlur</td>
<td>0.78</td>
</tr>
<tr>
<td>DomFp</td>
<td>0.86</td>
</tr>
<tr>
<td>ValidLinks</td>
<td>0.64</td>
</tr>
<tr>
<td>NumAds</td>
<td>0.70</td>
</tr>
<tr>
<td>Plagiarism</td>
<td>0.65</td>
</tr>
</tbody>
</table>
## Social Spam Detection

<table>
<thead>
<tr>
<th>Weka Classifier</th>
<th>Accuracy</th>
<th>FP</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LogitBoost</td>
<td>97.91%</td>
<td>.018</td>
<td>.978</td>
</tr>
<tr>
<td>LWL</td>
<td>97.68%</td>
<td>.013</td>
<td>.975</td>
</tr>
<tr>
<td>AdaBoostM1</td>
<td>97.68%</td>
<td>.018</td>
<td>.975</td>
</tr>
<tr>
<td>ConjunctiveRule</td>
<td>97.68%</td>
<td>.013</td>
<td>.975</td>
</tr>
<tr>
<td>DecisionTable</td>
<td>97.68%</td>
<td>.018</td>
<td>.975</td>
</tr>
<tr>
<td>DecisionStump</td>
<td>97.68%</td>
<td>.013</td>
<td>.975</td>
</tr>
<tr>
<td>RandomCommittee</td>
<td>97.45%</td>
<td>.018</td>
<td>.973</td>
</tr>
<tr>
<td>RandomForest</td>
<td>97.45%</td>
<td>.018</td>
<td>.973</td>
</tr>
<tr>
<td>Bagging</td>
<td>97.22%</td>
<td>.022</td>
<td>.970</td>
</tr>
<tr>
<td>NNge</td>
<td>97.22%</td>
<td>.022</td>
<td>.970</td>
</tr>
<tr>
<td>Features</td>
<td>linear SVM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>FP</td>
<td>$F_1$</td>
</tr>
<tr>
<td>TagSpam</td>
<td>95.82%</td>
<td>.061</td>
<td>.957</td>
</tr>
<tr>
<td>+ TagBlur</td>
<td>96.75%</td>
<td>.048</td>
<td>.966</td>
</tr>
<tr>
<td>+ DomFp</td>
<td>96.75%</td>
<td>.048</td>
<td>.966</td>
</tr>
<tr>
<td>+ ValidLinks</td>
<td>96.52%</td>
<td>.048</td>
<td>.964</td>
</tr>
<tr>
<td>+ NumAds</td>
<td>96.52%</td>
<td>.048</td>
<td>.964</td>
</tr>
<tr>
<td>+ Plagiarism</td>
<td>96.75%</td>
<td>.048</td>
<td>.966</td>
</tr>
</tbody>
</table>
Social Spam Detection

The diagram shows the percent correctly classified as a function of the number of features. The graph compares the performance of linear SVM and AdaBoost algorithms. The percent correctly classified increases with the number of features for both algorithms, with AdaBoost showing a slight advantage over linear SVM.
Related Work

- Web/Email Spam
  - Attenberg and Suel 2008, Gyöngyi et al. 2004
- Social Spam
  - Heymann et al. 2007
- Spam Detection
- ECML PKDD Discovery Challenge 2008
  - Held by BibSonomy team
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