

The Anti-Social Tagger - Detecting Spam in Social Bookmarking Systems

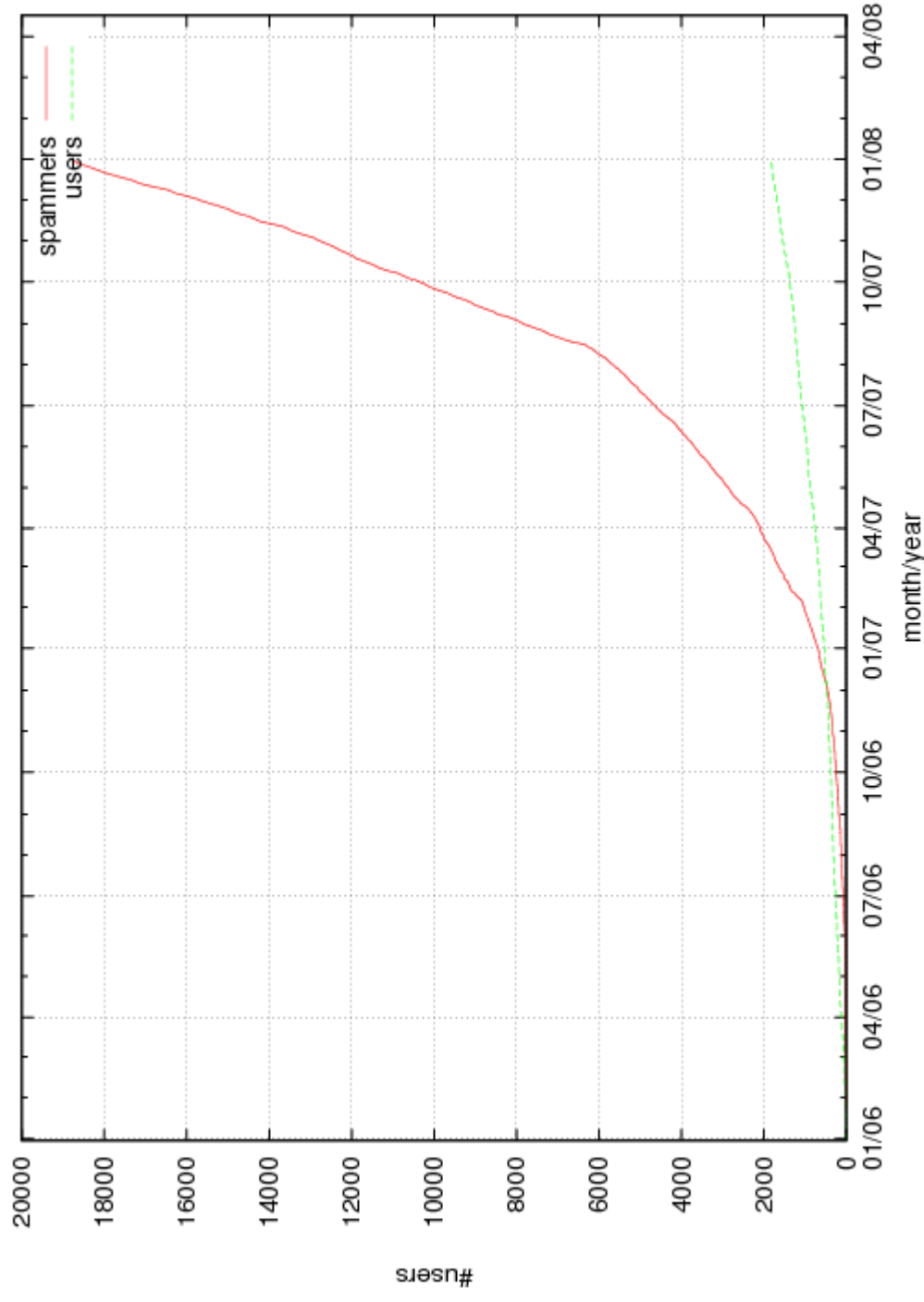
AIRWeb 2008

Beate Krause, Christoph Schmitz, Andreas Hotho, Gerd Stumme
Knowledge & Data Engineering Group,
University of Kassel &
Research Center L3S

AIRWeb 2008: The Anti-Social Tagger - Detecting Spam in Social Bookmarking Systems

3

BibSonomy “active” user accounts over time ...

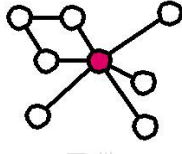


Problem Statement



Spam Detection in Social Bookmarking Systems

- What is social bookmarking spam?
- What kinds of features are there to detect spam?
- Can standard classification algorithms be used?



Agenda

■ Introduction

■ **Social bookmarking basics**

■ Spam definition

■ Dataset creation

■ Framework

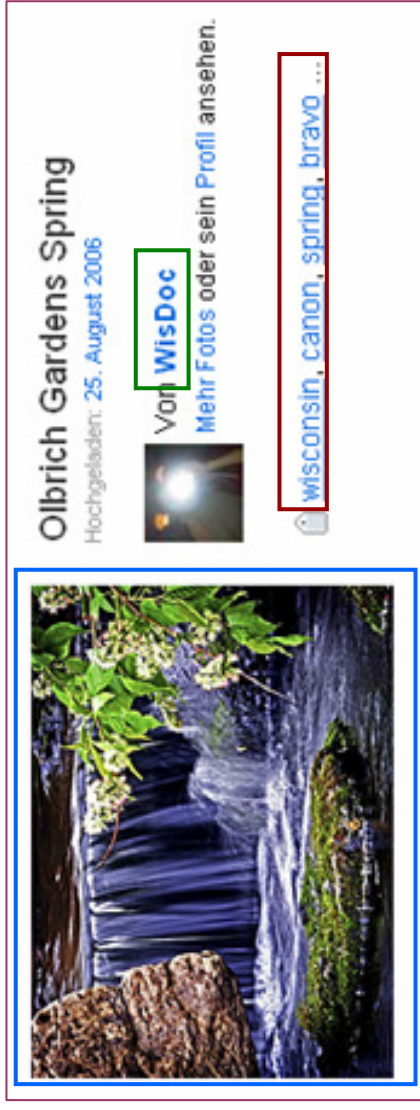
■ Evaluation

■ Conclusion

Folksonomy: Definition



- Folk + Taxonomy
- Folksonomies allow users to assign arbitrary tags to resources



- Tuple $F := (U, T, R, Y)$ where
 - U , T and R are finite sets, whose elements are called user, tags and resources
 - Y is a ternary relation between them, i.e. $Y \subseteq U \times T \times R$ whose instances are called **tag assignments (TAS)**,
 - A **post** contains a user, different tags and a resource.

[Andreas Hotho, Robert Jäschke, Christoph Schmitz, Gerd Stumme. Information Retrieval in Folksonomies: Search and Ranking. 2006]

Folksonomy: Social aspects



Search results for **spring**

Related tags: [spring java springframework framework tutorial j2ee programming hibernate developme](#)
showing 1 - 10 of 80561

[← previous](#) | [next →](#)

- [Springframework.org](#) save this
to java spring framework j2ee programming ... saved by 2059 people
- [Spring MVC step-by-step](#) save this
to spring java tutorial mvc framework ... saved by 804 people
- [CSS Reboot Spring 2006](#) save this
to css webdesign design inspiration gallery ... saved by 592 people
- [Enterprise Java Community: Introduction to the Spring Framework](#) save this
to spring java framework j2ee programming ... saved by 795 people
- [Java Certification Community: free Java J2ee scjp spring hibernate mock exams](#)
to java certification programming j2ee community ... saved by 1334 people
- [Acegi Security System for Spring - Acegi Security System for Spring](#) save this
to security java spring acegi authentication ... saved by 557 people
- [UC Berkeley Webcasts | Video and Podcasts: Spring 2007 Courses](#) save this
to education podcast learning free webcast ... saved by 1507 people
- [Enterprise Java Community: Introduction to the Spring Framework 2.5](#) save this
to spring java framework j2ee programming ... saved by 578 people
- [Acegi Security System for Spring - Acegi Security System for Spring](#) save this
to security spring java framework authentication ... saved by 529 people
- [Enterprise Java Community: Introduction to the Spring Framework](#) save this
to spring java framework j2ee programming ... saved by 452 people

[← previous](#) | [next →](#)

- Easy to use, open for everyone
- Collaborative annotation of web resources
- Immediate “feedback”
- Joint use leads to converging vocabularies and a collection of web resources



Agenda

■ Introduction

■ Social bookmarking basics

■ **Spam definition**

■ Dataset creation

■ Framework

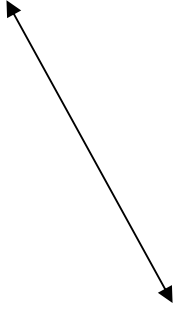
■ Evaluation

■ Conclusion

Spam: Malicious Tagging - vs. Web Spam



Website



Post

Net4 Budget Hotel Guide Peking Peking/Beijing

to clustering community-detection web-mining by bsb2000 on Apr

17, 2008, 9:34 AM

Net4 Hotelguide: Budget Hotels und Pensionen in Peking
Budget Hotels - das sind preisgünstige Hotels im Niedrigpreis-Bereich.

In [Peking](#) gibt es Tausende
Übernachtungsmöglichkeiten.
Bei Auswahlmöglichkeiten ist nicht
schwer zu entscheiden. Hier
sind einige Tipps, die Ihnen
helfen können.

Anzeige
[China Visum Chinesisches](#)
[Visum Chinesische](#)
[Botschaft](#)
[Eastline-Visa-Service:](#)
[Touristen- und](#)
[Geschäftsvisa ab 6 STD](#)

mehreren deutsch- oder
englischsprachigen
Reiseführern empfohlen
wurden.

1 Euro = 9.67277813 Yuan
am 16.3.2006, nach google
search "Euro in Yuan".

Preisangaben nur ungefähr.
Liste erstellt 2001. Update
2005

[Qiaoyuan Hotel](#)

Anzeige
Das richtige Hotel finden
in China mit [Booking.de](#)

Beijing Capital Flughafen
Hotels in der Nähe der Flughafen. Schnell
und sicher online buchen
20 Hotels in Beijing
Bequem und Schnell online buchen!
Beijing Stadtplan, Fotos
Russland Visum & Einladung
Visum ab 55 € inkl. Einladung, Visum in 3
Stunden Tel: 02232566157
Zhaolong Peking Beijing
Hotels finden, Preise vergleichen, Kritiken
lesen TripAdvisor testen!
AIV Google-Anzeigen

Spam - User Level vs. Post Level



BibSonomy :: user ▾ :: rygar33 ::

A blue social bookmark and publication sharing system.

[tags](#) · [relations](#) · [groups](#) · [popular](#)
[myBibSonomy](#) · [post bookmark](#) · [post bibtex](#)

bookmarks (4) [RSS](#) [XML](#)

<< < 1 > >>

[edit](#)

Home Equity Line of Credit

Heloc options for your credit type. JumpStartMyCredit.com

to 2nd cash credit equity free heloc helocs home line loan loans mortgage of out quote rate second by

rygar33 and 4 other people on Apr 17, 2008, 9:12 AM

[copy](#)

Home Loan Refinance

Mortgage Refinance Quotes. No SSN Needed.

to calculator cash compair compare home in loan lock lower mortgage out payment quote rate refi

refinance refy by rygar33 and 4 other people on Apr 17, 2008, 6:49 AM

[copy](#)

Mortgage Quote

Home Loan Quotes. No SSN Needed.

to bad check credit home interest lenders lendingtree loan mortgage no purchase quote rates by

rygar33 and 4 other people on Apr 17, 2008, 3:39 AM

[copy](#)

Online Loan Source for Bad Credit

Payday Loans, Home Loans, Bad Credit Auto Loans, Refinance, Debt Settlement Services.

to auto bad car consolidation credit financing loan loans payday student by rygar33 and 4 other people

on Apr 17, 2008, 1:28 AM

[copy](#)

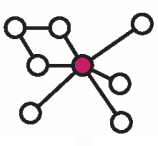
<< < 1 > >>

bookmarks per page: 5 10 20 50 100

Spam in this work



- Web Spam:
 - Interaction with the system that serves commercial, self-promoting or disruptional interests
 - Interactions that are not of value for other users
- Target for classification: user, not post



Agenda

Introduction

Social bookmarking basics

Spam definition

Dataset creation

Framework

Evaluation

Conclusion

Dataset



- BibSonomy admins and developers flag users as spammers
- Decision is based on
 - Links (websites) of posts
 - Added tags
 - Also influenced by personal information:
 - E-mail
 - Choice of name
 - IP
 - ...



3

Dataset creation process



user :: adomorie6

Mozilla Firefox

datei

bearbeiten

ansicht

chronik

lesezeichen

extras

hilfe

postBookmark

postPublication

SpammerKillerButton

www.lego.org

postBookmark

BibSonomy :: user :: adomorie6

http://www.bibsonomy.org/user/adomorie6

Unbenannt

user :: adomorie6

tags · relations · groups · popular

myBibSonomy · post bookmark · post bibtex

bookmarks (31)

rss

xml

publications

publications per page: 5 10 20 50 100

logged in as beate · help · blog · about

22 picked in basket · edit tags · settings · logout

filter:

- relations

- tags

(alpha | freq) (cloud | list) (minfreq 1 | 2 | 5)

"credit \ \ \ \ accept advisor alternative amex amygoodman Annapolis annapolis Anthem approved assad atkins attitude auto autobody autoinstruction automotive autoparts bad bank banks biodiesel bloodpressure boys budget Buffet bus bush california campus Canada car carbohydrates carbs card" Career cars cash cater caterer chance cheney Chevy cholesterol.chelestorel cia college Conaleeza condaleza condi condirice condo conflict contacts cottage courses credit danger dealer dealership debt demerit democrats detroit diabetes diesel diet dieting diets domestic ecologist ecology employment energy estate ethanol fat fats Fenway finance financial financing Fish Fishing food Ford franchise g.m. G.M.C.. game garage gas gasohol golan golf golf.golfing

Star Wars: LEGO's Ultimate Collectors Millennium Falcon Unboxed, Now Shipping with 4-pound Manual - Gizmodo Ultimate Lego Set

to boys lego toys by adomorie6 and 63 other people on Apr 21, 2008, 1:39 AM

copy

Sarbit Asset Management

Our Primary Concern is to Ensure that Client's are Protected Everything else follows from that fundamental concept in investing

to Buffet Larry Sarbit Warren portfolio sarbit by adomorie6 and 66 other people on Apr 21, 2008, 1:38 AM

copy

Are You Becoming Wealthy On Your House?

Are you becoming wealthy on your house... Is your home your best performing investment...

to \ \ \ auto bad car finance home house invest investment mortgage realtor by adomorie6 and 67 other people on Apr 21, 2008, 1:36 AM

copy

Fertig

AIRWeb 2008: The Anti-Social Tagger - Detecting Spam in Social Bookmarking Systems

16

Data Creation Process



Are You Becoming Wealthy On Your House? - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.isnare.com/?aid=186991&c=Finances

postBookmark postPublication SpamKillerButton www/leo.org postBookmark

Are You Becoming Wealthy On Yo... (Unbenannt)

Free content articles for your ezine and website.

Home - Register - My Account - Advertise - Link to Us - FAQs - Site Map

Titles Authors Contents [Advanced Search][Add OpenSearch][Job Search]

Isnare.com Articles

Distribute your articles to thousands of article sites for only \$2 and below! Read More >>

Search

To reprint, publish or integrate this article in any website, publication or by any means read and follow our [Acceptable Use Policy](#).

[The Real Estate Market](#)
Get news and expert opinions from The New York Times realstate.nytimes.com

[We pay high interest rate](#)
Up to 14% daily. Fast and Safe. Online Investment Service www.top1profit.info

[Free China Stock Report](#)
3 Safe and Easy Ways for US Investors To Own Chinese "T-Shares"

Ads by Google

[Contact the Author] [Send to a Friend] [Article Publisher] [Make PDF] [Print] [Bookmark & Share]

Are You Becoming Wealthy On Your House?

Submitted By: [Kerry A. Francis](#)

[Word Count: 978 words] [Viewed: 2620 time(s)] [Don't forget to rate this article] [Report this article]

Are you becoming wealthy on your house? Is your home your best performing investment? Is your house the only area that of your investments in which you are making money?

Red danger signals should be appearing in your mind. The housing market has gone up, up and up. Many people believe that they "have made x dollars

[S&P 500 Trading Signals](#)
100% Success From Our System
Trade Our Signals And Make Money
www.stockmis.com

[Investment Opportunity.](#)
Big Potential Returns On Investment
Apply Online For More Information.
StrategicInvestmentsUK.com

[Do you want in on China?](#)
We have a list of companies that will make it HUGE in China. Free.

Fertig

Dataset creation process



BibSonomy :: user ▾ :: adomorrie6 ::

A blue social bookmark and publication sharing system.

Dataset creation process



BitSonomy:admin - Mozilla Firefox

postBookmark postPublication SpammerKillerButton WWW leo.org postBookmark

BitSonomy:admin

Info: user 'adomorie6' flagged as spammer!

flag/unflag a spammer

Spammer: user name

Spammer: user name

User Info

user name

no log messages ...

Scrapper

Highwire Linklist

Last Update: 09.07.2007 13:26:22

API key

please select a user

add a group to the system

user name

privacy

existing groups

abcdi

aicud

aksw

apostle

bibliothek2.0

cais

cambio

care

cdg

cisse

ckc2007

cscw

css739

daks

dame

digpio

div3

dta

dti

echo

eni

erealtat_ibk_07

evol

hist3718

hpi.its

iccs

iim_mags_socsoft

imo

imtm

ip

iswc2007

List of new users

(show 100 500 1000 all, show frequent IPs)

Is it a Spammer?	Username	Spammer?	IP	Realname	E-Mail	Date	updated by	updated at
YES NO	chuck37	false				21.4.08_01:17		1815-12-10 00:00:00.0
YES NO	mourad	false				21.4.08_01:07		1815-12-10 00:00:00.0
YES NO	adarinda	false				21.4.08_00:58		1815-12-10 00:00:00.0
YES NO	argilla_educator	false				21.4.08_00:41		1815-12-10 00:00:00.0
YES NO	eusem1956	false				21.4.08_00:04		1815-12-10 00:00:00.0
YES NO	berry48	false				21.4.08_00:03		1815-12-10 00:00:00.0
YES NO	ponoro	false				20.4.08_23:58		1815-12-10 00:00:00.0
YES NO	zostad	false				20.4.08_23:51		1815-12-10 00:00:00.0
YES NO	adomorie6	true				20.4.08_23:50	beate	2008.04.21 01:46:22.0
YES NO	ppockik	false				20.4.08_23:49		1815-12-10 00:00:00.0
YES NO	uftrust_user	false				20.4.08_23:38		1815-12-10 00:00:00.0
YES NO	gritteeth	false				20.4.08_23:16		1815-12-10 00:00:00.0
YES NO	alighisinto	false				20.4.08_22:36		1815-12-10 00:00:00.0
YES NO	musicmp3	true				20.4.08_22:33	jaeschke	2008.04.20 23:15:51.0
YES NO	pasley77	true				20.4.08_21:50	jaeschke	2008.04.20 00:00:00.0

Dataset creation process



list of spammertags (652)

most used spammertags (last 1000 posts)

tag	count	action
and	75	↗
for	54	↗
sex	29	↗
dildo	23	↗
the	22	↗
of	22	↗
sale	20	↗
Money	20	↗
anal	19	↗
a	19	↗
butt	19	↗
vacation	16	↗
to	15	↗
in	15	↗
deals	15	↗
health	14	↗
gold	12	↗
ass	12	↗

tag suggestions

tags used by spammers but not marked as spammertags (last 1000 tas)

tag	count	action
download	70	↗
music	67	↗
mp3	65	↗
downloads	64	↗
internet	52	↗
meta	47	↗
Engineering	47	↗
search	46	↗
Business	46	↗
Ingenieur	45	↗
Kartographie	44	↗
Bergbau	44	↗
Jobs	44	↗
video	43	↗
Architektur	42	↗
software	41	↗
Music	41	↗
tool	40	↗

spammer suggestions

users not flagged as spammers but used spammertags

user	tagged action
rorygearhart	↗ of ↗
franz20	↗ Free ↗
msswongi	↗ in ↗
kingtom	↗ cheap ↗
mabelca29	↗ Free ↗
jenniferbonaccorsi	↗ money ↗
greenic1980	↗ health ↗
sundhed.helbred	↗ health ↗

Dataset - Figures



	Users	Spammer	Tags	Posts	TAS
All	1,406	18,681	306,993	1,219,053	8,709,417
Training	1,306	15,891	282,473	1,097,458	7,904,735
Test	100	2,790	49,644	121,595	804,682

- Time frame: until end of 2007
- Only users with at least one post
- No consideration of private posts
- Not normalized

Dataset Issues



- Subjective notion of what is spam
- No crosschecks of flagging results → noise
- Only two classes: spam and non-spam
- User level not granular enough
- One user may have several accounts



Agenda

■ Introduction

■ Social bookmarking basics

■ Spam definition

■ Dataset creation

■ **Framework**

■ Evaluation

■ Conclusion

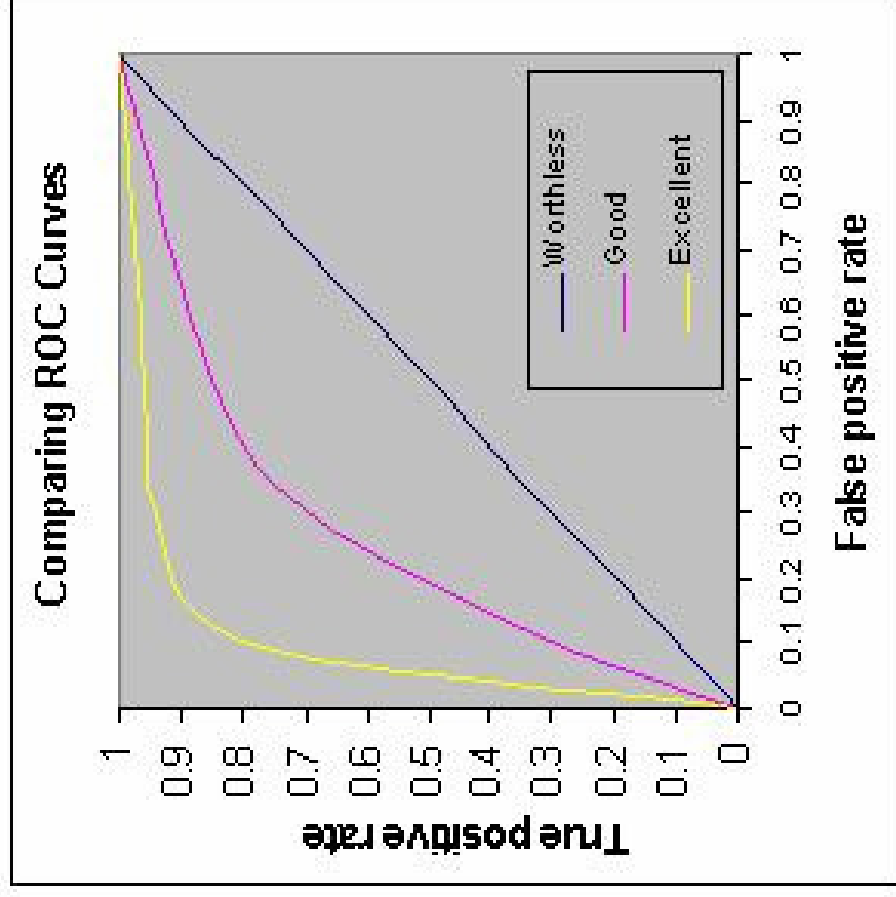
Evaluation measures



Setting

Actual/ Labeled	Spam	Non-Spam
Spam	TP	FN
Non-Spam	FP	TN

- Precision, Recall, F1
- ROC Area

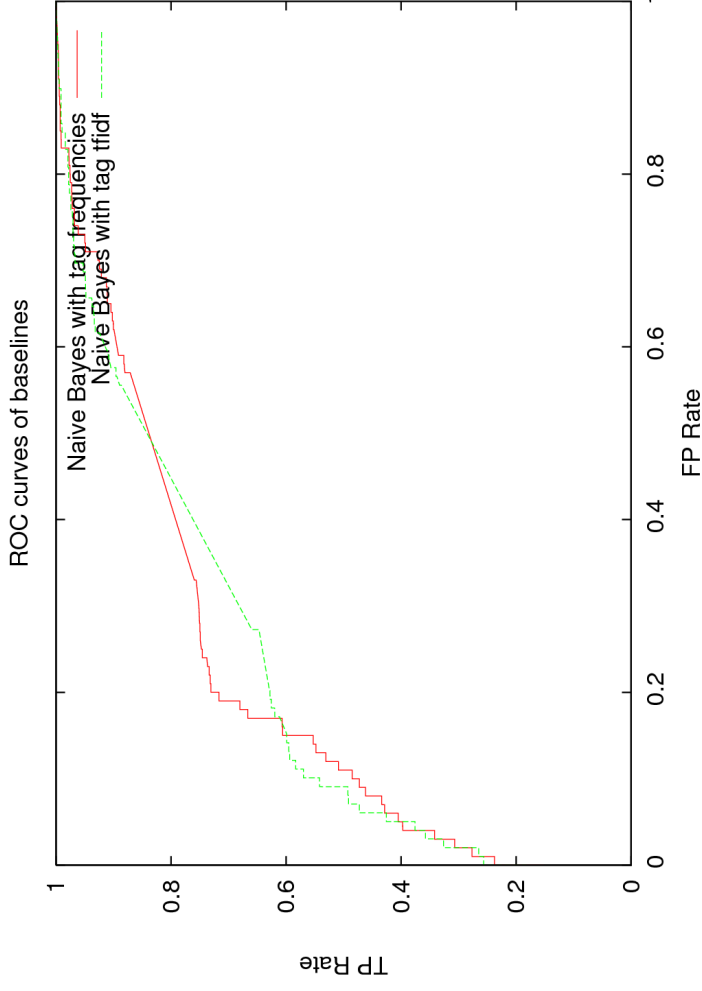


[<http://gim.unmc.edu/dxtests/ROC3.htm>]

Baseline



- Features: different tags
- Instance (one user): represented by the tags added to all of his/her posts
- Either frequency, or tfidf values per tag
- Naïve Bayes classification



Baseline



- Frequency ROC Area: 0.80
- TFIDF ROC Area: 0.79

Table 4: Baseline with all tags as features (frequency)

	Spam	Non-Spam
Spam	466	2324
Non-Spam	0	100

Table 5: Baseline with all tags as features (tfidf)

	Spam	Non-Spam
Spam	530	2260
Non-Spam	0	100

Features



- 25 features
- four different categories
- normalization of each user's feature vector

Profile Information

Activity Information

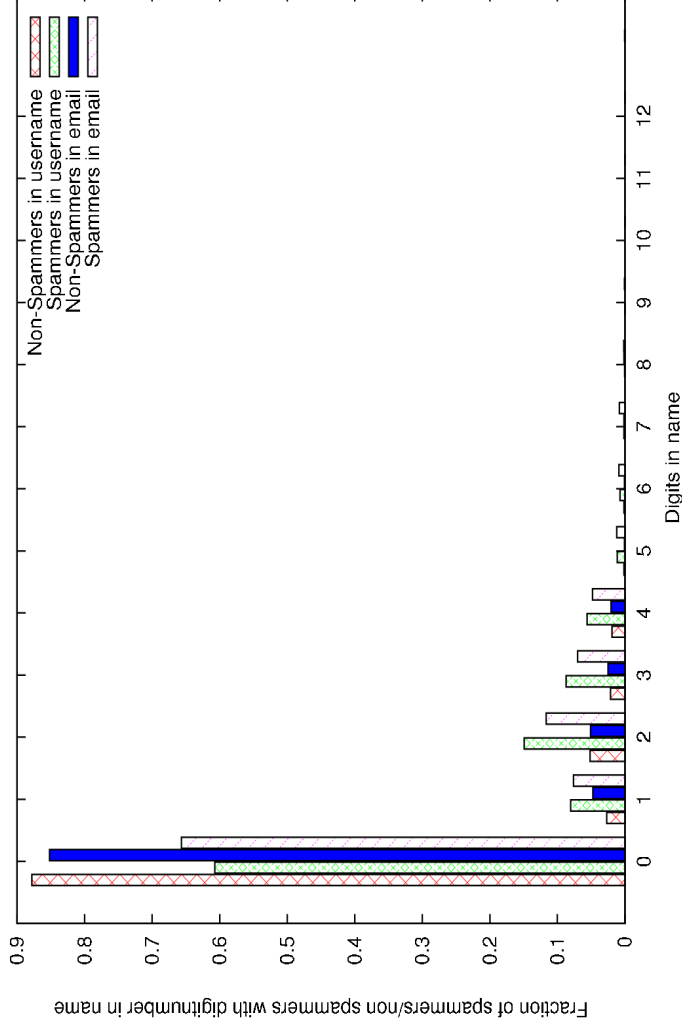
Location Information

Semantic Information

Profile features



- *namedigit*: „styris888“, „painrelief2“
- *maildigit*: digits in mail
- *namelen*, *maillen*, *realnamelen*: length of names, mails
- *realnameN* (2,3): binary, set to 1 if user has indicated two or more names



Location based features



- *domaincount*: number of users in the same domain
 - 16 users: spambob.com
 - 137 rhinowebmail
- *tldcount*: number of users in the same top level domain
- *spamip*: number of spam users with this IP
 - one spammer uses several accounts
 - accounts often have the same IP
- [future work: what part of the world does IP come from?]

Activity based features



- *datediff*: Time between registration and first post
- *tasperpost*: number of tags per post
 - Spammers: 8, Users 4
- *tascount*: average number of TAS
 - 470 for spammers, 334 for users

Semantic features



- *\$Group=public*: automatic tag from spamming software
- *spamtag*: blacklist of tags

tag	count	action		vacation	
and	75			16	
for	54			15	
sex	29			15	
dildo	23			15	
the	22			14	
of	22			12	
sale	22			12	
Money	20			bodybuilding	
anal	20			11	
a	19			car	
butt	19			10	
vacation	16			poker	
to	15			gift	
in	15			8	
deals	15			earrings	
health	14			diamond	
gold	12			shirt	
ass	12			black	
				money	
				loans	
				diet	
				loan	

Semantic features: cooccurrences



- *Cooccurrence information:*
 - Undirected graph with user as nodes
 - Link exists if users share at least one resource, one tag or one tag-resource pair
 - Each case twice: spammer / non spammer
 - Example: A spammer shares resources with about 18 other spammers, but only with 0.5 non-spammers
- *Ratios of Spam / Non-spam pairs*



Agenda

Introduction

Social bookmarking basics

Spam definition

Dataset creation

Framework

Evaluation

Conclusion

Evaluation of all features



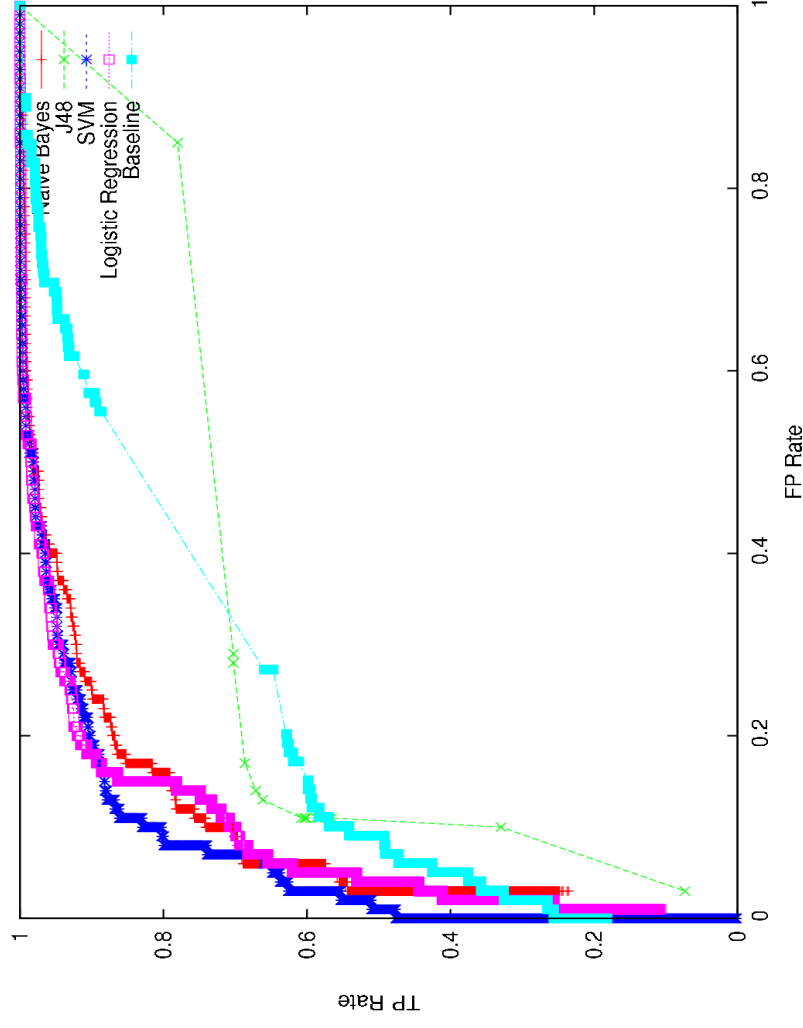
Tool:

- Weka

Classification Algorithms:

- SVM
- J48
- Logistic Regression
- Naïve Bayes

All features



Evaluation of all features (+costs)



Table 10: Evaluation values all features

Classifier	ROC area	F1	FP	FN
Naive Bayes	0.906	0.876	14	603
SVM	0.936	0.986	53	23
Logistic Regression	0.918	0.968	30	144
J48	0.692	0.749	11	1112

Table 12: Evaluation with a cost sensitive classifier

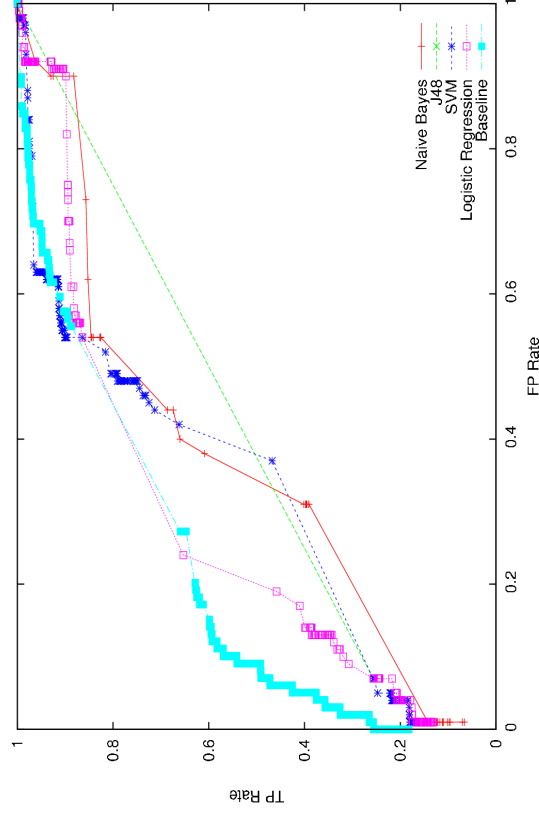
Classifier	F1	FP
SVM	0.924	15
J48	0.794	11
Logistic Regression	0.927	12
Naive Bayes	0.855	11

- Best classification algorithm: SVM
- But: SVM has highest FP Rate! → Cost sensitive classification
- AUC values of baseline yield lower results

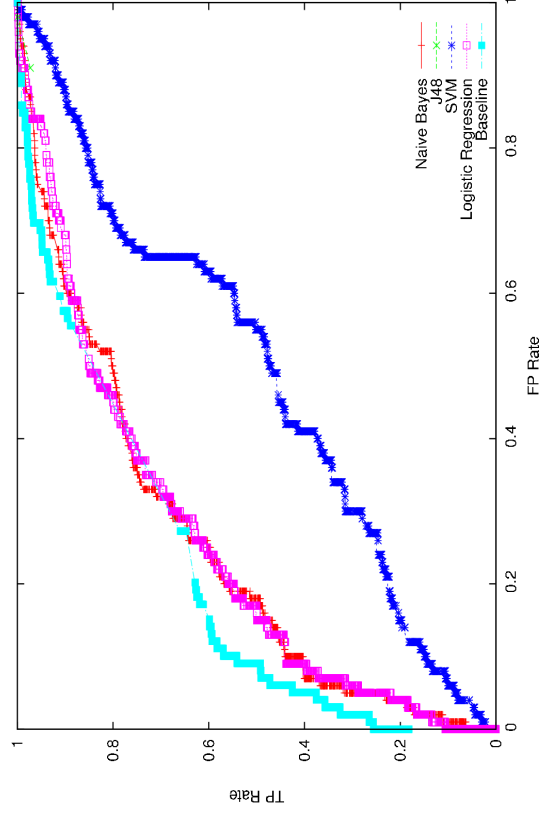
Evaluation of feature groups



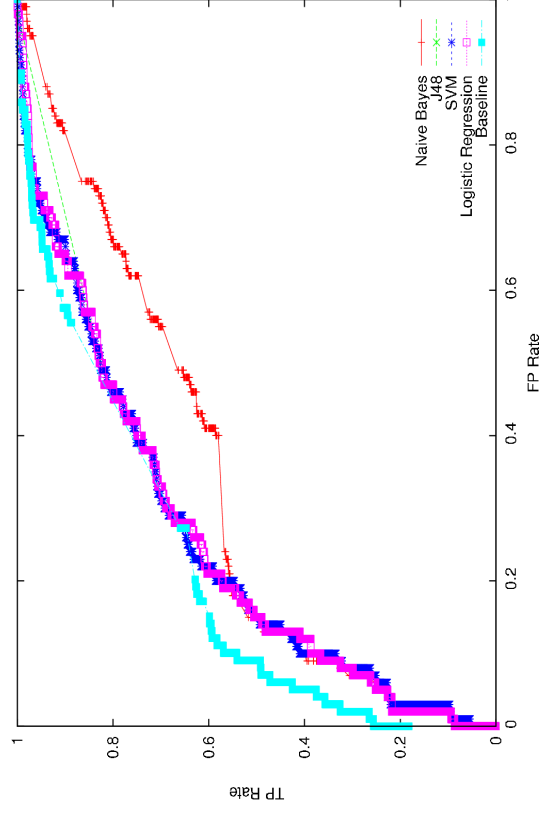
Location based features



Profile based features



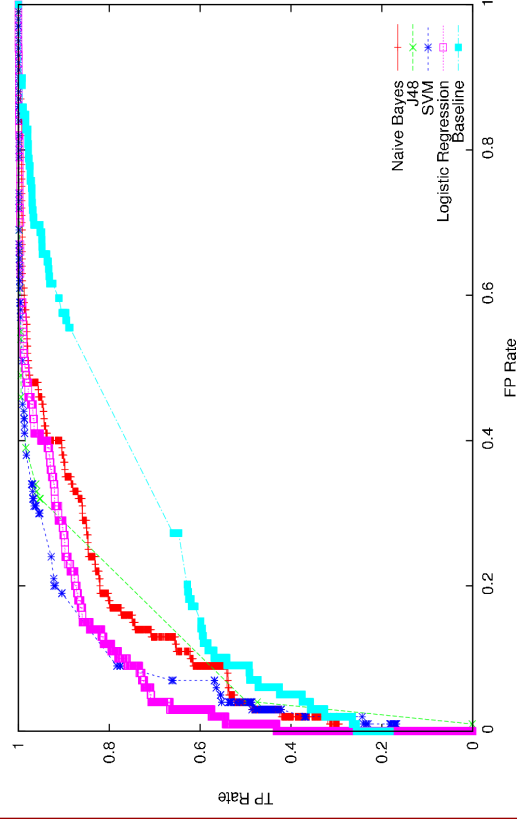
Activity based features



Evaluation all features



Semantic features (cooccurrence)



Semantic features (group tag, spam tag)

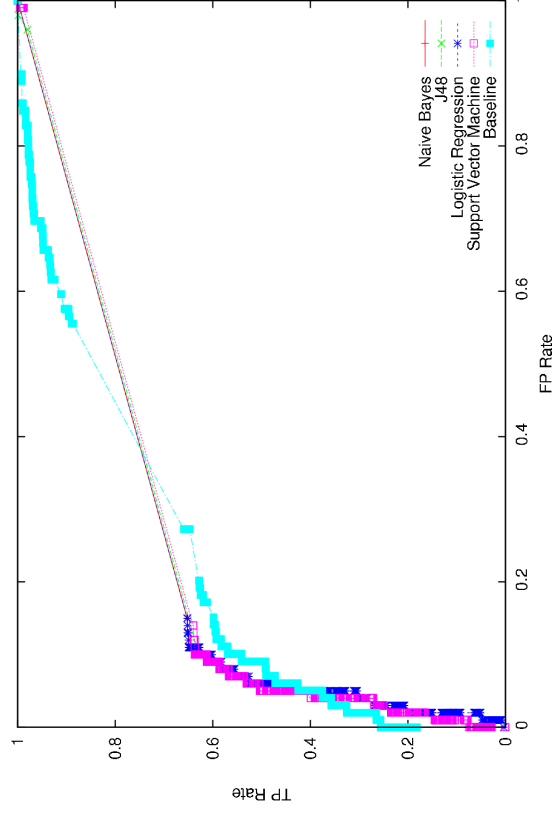


Table 11: Evaluation values of feature groups

Features	ROC area	F1
Profile features (log. reg.)	0.77	0.982
Location features (SVM)	0.698	0.407
Activity features (SVM)	0.752	0.982
Semantic features (J48)	0.815	0.981
Cooccurrence features (log. reg.)	0.927	0.985



Agenda

Introduction

Social bookmarking basics

Spam definition

Dataset creation

Framework

Evaluation

Conclusion

Conclusion



- Many features from users based information
- Combining all features: tangles problem quite well
- SVM best, but needs cost-sensitive classification
- Different feature groups: Co-occurrence features most promising, but computation intense

Future work



- Fine-tuning of results
- Reduce user accounts
- Classification on post level
- Use different splits
- Add more features
- Comparison with other techniques (e.g. language models)
- Comparison with other datasets



Agenda

Introduction

Social bookmarking basics

Spam definition

Dataset creation

Framework

Evaluation

Conclusion