DIG Student Seminar

Spamming on the Social Web

Albert Au Yeung

12 November 2008



- Problem of spamming on social Web sites
- Types of anti-spamming strategies
- Evaluation of anti-spamming techniques
- Spamming in collaborative tagging



- Social Web sites: content is primarily supplied by general users
- New means of sharing/discovering useful/interesting resources on the Web
- Also interested by spammers: new means of attracting user attentions
- Spam hinders resource sharing and discovery (e.g. 19 of 20 top most active users in Delicious are spammers (Wetzker et al. 2008))



Characteristics of Social Web Sites

• Centralisation:

one controlling entity, runs its own servers, has access to all the data

• Well-defined interactions: user behaviours tend to be more predictable

- Identity: every user's action can be traced
- Multiple interfaces: more chances to get spammed, more different anti-spamming techniques are required



Example: Delicious

\sim	Popular design Bookmarks Recent Popular			Look up a URL	
	See all design bookmarks.				
Popular	design x			Tags	
				* Related Tags	1
	Newspaper Website Design: Trends And Examples I Design Show	case Smashing	305	inspiration	
	Magazine save	date for and		webdesign	
	First seven by colementer informatition cool we	absite userexperience	resource	art	
	Fischeres of Costinuing			web	
	18 Rules the Best Web Developers Follow Design And Marketing	BIOG SAVE	229	portfolio	
	First saved by: stuartgoff	list tips wel	design	blog	
	17 Productive Photochan Cheatchaste and Poferance Carde to De	welload for Eroo :	327	illustration	
	Speckyboy - Web Design Web Development and Graphic Design	Resources save		graphic	
	opeckyboy - web besign, web bevelopment and oraphic besign	ne tutorial resources	design	typography	
	First saved by Tedegrato		A accedent	resources	
	50 Must-Have Photoshop Brushes Developer's Toolbox Smashir	ng Magazine SAVE	987		
	First saved by: LJK brushes	photoshop tools free	e design		
			497		
	45 Beautiful 3D Typography Designs For Inspiration SAVE				
	First saved by, mindxploder		design		
	Design Patterns I Object Oriented Design SAVE		543		
		patterns java pattern	design		
	First saved by, Inchi_sun				



Browsing bookmarks with a certain tag

Browsing popular bookmarks





Detection-based Anti-spamming

- Simplest form: manual identification
- Pattern-based automatic identification
- Usually involves supervised learning (need training data or seed data)
- Example in tagging: common patterns include (1) bookmarks within the same domain (2) use title words as tags
 - (3) use same tags across many bookmarks

Anti-spamming Strategies

	dinojacob's Bookmarks Bookmarks Network Tags Subs	criptions	
-	See more bookmarks in Popular or Recent.		
nojaco	ob Type a tag	Bookmarks 61 Tags	
= =		Sorted by Most Recent 🔻 🔻 Top 10 Tags	
IOV 08	Bad Credit Refinance SAVE	refinance	20
		mortgage (refinance, (loan (calculator, (bad (credit (mortgage, loan	20
		refinance (interest (rates, (home (equity (loans, (best	15
	Bad Credit Mortgage SAVE	mortgage	15
	Dad Orean Mongage 3AVE	mortrage refinance loss calculator bad credit mortrage	15
		refinance interest rates home equity have best	14
		nome equity iouns, bost nome equity iouns, bost nome	14
	Best Refinance SAVE	equity	14
		mortgage refinance, loan calculator, bad credit mortgage, refinance.	14
		refinance interest rates, home equity loans, best	42
	Bad Credit Mortgage Refinance	SAVE	42
		mortgage < refinance, < loan < calculator, < bad < credit < mortgage,	
		refinance (interest (rates, (home (equity (loans, (best	
	Home Loan Home Loan Refina	DCA SAVE	
	Home Edul (Home Edul (Keine	mortgage refinance loan calculator bad credit mortgage	
		refinance interest rates, home equity loans, best	
	Mortgago Pofinancing		
	Mongage Reinlanding SAVE		
		mortgage refinance, loan calculator, bao credit mortgage,	
		remnance interest rates, nome equity loans, best	
	melinablack.sosblog.com - Blog	The first blog : Last posts SAVE	
		mortgage refinance, loan calculator, bad credit mortgage,	
		refinance interest rates, home equity loans, best	
	Mortgage Refinance Auto Loar	Auto Loan Refinance Bad Credit Auto Loan	

Demotion-based Anti-spamming

- Reduce the prominence of spams (remove it from the top of the lists)
- Ideally providing ranking which promotes good content and demotes spams at the same time
- With a good algorithm, this should be the most desirable form of anti-spamming
- Need different algorithms for different interfaces (e.g. most recent, most popular, single user/tag)

Prevention-based Anti-spamming

- Making spamming activities more difficult to carry out
- Example: using CAPTCHA



- Make contributing content cost computational time or even real money (e.g. the PennyBlack project at Microsoft)
- Problem: causing inconvenience to real users at the same time

Spam Models

- Give a model of the spamming activities on a site
- Describe what a spam would look like
- (1) synthetic model
 - representation of a social system
 - definition of spamming (e.g. wrong tags)
 - may not capture activities of real spammers
- (2) trace-driven model
 - require labelling of real data
 - produce more realistic results
 - but may result in bias, also time-consuming

Spam Metrics

- Quantify the impact of spam on an interface
- Measure effectiveness of a anti-spamming technique
- Examples:
 - (1) Precision & Recall
 - (2) SpamFactor (Koutrika et al. 2007)

$$SpamFactor(n,t) = \frac{\sum_{i=1}^{n} w(o_i,t) * \frac{1}{i}}{\sum_{i=1}^{n} \frac{1}{i}},$$

where
$$w(o_i,t) = \begin{cases} 1 \text{ if } t \text{ is a bad tag for } o_i \\ 0 \text{ if } t \text{ is a good tag for } o_i \end{cases}.$$

Types of spammers

- Flooders
 - Tag a large number of existing documents
 - Aim at gaining reputation/high ranks

• Promoters

- Contribute a lot of new documents (spam Websites)
- Aim at promoting their own documents
- Trojans
 - Disguise themselves by contributing popular bookmarks
 - Direct attention of users to malicious Web sites

Spamming in Collaborative Tagging

Bookmarks (Network Tags Subscriptions See marks bookmarks in Popular or Recent. Indyrocotiman Type a tag Sorted by Most Recent * B oCT 08 forex currency trading beginner sAVE Currency trading system SAVE 12 Currency trading system SAVE 12 Currency forex online trading save 12 Currency trading save 11 Currency day trading save 11 Currency trading system save 12 Currency trading save 12 Torex, currency, trading, currency trading system 12 Currency trading save 12 Trading, currency trading save 12 Torex, currency, trading, online currency trading 12 Torex, currency, trading, currency forex trading 12 Torex, currency, trading,	P	judygoodman's Bookmarks	 Save a new bookma Tag options 	ark
see more bookmarks in roppuar or record. judygoodmain judygoodmain Type a tag Sorted by Most Recent a oct rols forex. currency, trading, beginner, forex, trading, currency, trading, beginner, forex, trading, currency, trading, system, currency, beginner, currency, trading, system, currency, trading, foreign, currency, trading, save forex, currency, trading, save 12 instrance 20 instrance 21 forex, currency, trading, savte <th></th> <th>Bookmarks Network Tags Subscriptions</th> <th></th> <th></th>		Bookmarks Network Tags Subscriptions		
Improvidential programmer Improv	iudugoo	See more bookmarks in Popular or Recent.	Terre	
a octros forex currency trading beginner save 12 forex, currency, trading, beginner, forex forex, currency, trading, currency forex, currency, trading, beginner, forex forex, currency currency trading system save 12 currency forex online trading save 12 currency, forex, online, trading, currency online forex 11 e currency, trading, save 12 currency trading save 11 e currency, trading, currency day currency trading forex, currency, trading, currency trading forex, currency, trading, save 12 currency trading save 12 foreign currency trading save 12 online, currency, trading, online 12 forex, currency, trading, currency, trading, online 12 forex, currency, trading, currency, trading, online 12 foreign currency trading save 12 forex, currency, trading, currency, trading, online 12 forex, currency, trading, currency, trading, currency forex trading </td <td></td> <td>Sorted by Most Recent V</td> <td>Tags</td> <td></td>		Sorted by Most Recent V	Tags	
solid is follow currency trading beginner forex, currency, trading, beginner, forex trading currency forex, currency, trading, beginner, forex trading currency forex fore	0.007.00	forev surronsy trading beginner aug	+ Top to Tags	20
torex, currency, trading, beginner, torex, trading currency beginner beginner 12 currency trading system save 12 currency trading save 11 currency trading save 12 currency trading save 11 currency trading save 12 currency trading save 12 currency trading save 11 currency day trading save 12 currency trading save 12 currency trading save 12 currency trading save 12 forex, currency, trading, currency trading system 14 forex, currency, trading, system, forex currency trading 14 foreign, currency, trading, system, foreign, currency trading 14 foreign, currency, trading, system, foreign, currency trading 14 foreign, currency, trading, system 12 foreign, currency, trading, online currency trading 12 forex, currency, trading, currency, trading, currency trading 12 forex, currency, trading, currency, trading, currency trading 12 forex, currency, trading, currency, trading, currency trading 12 forex, currency, trading, currency, tra	8 OCT 08		insurance,	20
currency trading system save 12 currency trading system save 12 currency forex online trading save 12 currency forex online trading save 12 currency trading save 11 e currency trading save 11 e currency trading save 11 currency day trading save 12 currency trading save 12 currency trading save 11 forex currency trading save 12 foreign currency trading save 12 online currency trading save 12 forex currency trading save 12 online currency trading save 12 forex currency trading save 12 online, currency, trading, online currency trading 12 forex currency trading save 12 online, currency, trading, online currency trading 12 forex currency trading save 12 currency trading save 12 forex currency trading save 12 currency trading save 12 currency, trading, currency forex trading 12 currency, trading, currency forex trading 12 currency, trading, currency forex trading 12 currency, trading, currency trading 12 forex, cu		torex, currency, trading, beginner, torex trading currency	personal	10
currency trading system save 12 currency forex online trading save 12 currency, forex, online, trading, currency online forex trading 10 recovery, trading save 11 e currency trading save 11 e, currency, trading, e currency trading 10 currency day trading save 11 currency day trading save 12 currency day trading save 12 currency trading system save 11 forex, currency, trading, currency day currency trading 162 foreign currency trading save 11 foreign currency trading save 12 foreign currency trading save 12 foreign currency trading save 12 forex currency trading save 12 forex currency trading save 12 forex currency trading save 12 currency, trading, online currency trading 162		beginner	injury	10
currency, trading, system, currency trading system 12 currency forex online trading sAVE 12 currency trading sAVE 11 e currency trading sAVE 11 e, currency, trading, e currency trading 10 currency, day trading sAVE 12 currency, day, trading, currency day currency trading 162 forex currency, trading system SAVE 11 foreign currency trading system SAVE 11 foreign currency trading save 12 online, currency, trading, foreign currency trading 14 foreign, currency, trading, online currency trading 14 forex, currency, trading, online currency trading 14 foreign, currency, trading, online currency trading 12 foreign, currency, trading, online currency trading 12 forex, currency, trading, online currency trading 12 forex, currency, trading, online currency trading 12 forex, currency, trading, currency forex trading 12 currency, trading, currency forex trading 12 forex, currency, trading, currency forex trading 12 forex, currency, trading, currency forex trading 12 forex, currency, t		currency trading system SAVE 12	currency	10
currency forex online trading SAVE 12 currency, forex, online, trading, currency online forex trading 11 e currency trading SAVE 11 e, currency, trading, e currency trading 12 currency day trading SAVE 12 currency, day, trading, currency day currency trading 162 forex, currency, trading, system, forex currency trading 162 foreign currency trading SAVE 12 conline, currency, trading, foreign currency trading 162 forex, currency, trading, online currency trading 12 forex, currency, trading, online currency trading 12 forex, currency, trading, currency forex trading 12 currency trading SAVE 12 forex, currency, trading, currency forex trading 12 currency trading SAVE 12 forex, currency, trading, currency trading 12 <td></td> <td>currency, trading, system, currency trading system</td> <td>life</td> <td>10</td>		currency, trading, system, currency trading system	life	10
currency, forex, online, trading, currency online forex trading recovery, 100 e currency trading save 11 e, currency, trading, e currency trading 11 currency, day, trading, currency day currency trading 12 forex, currency, trading, system 11 forex, currency, trading, system, forex currency trading 11 foreign, currency, trading, system, forex currency trading 12 foreign, currency, trading, foreign currency trading 12 online, currency, trading, online currency trading 12 forex currency trading save 12 online, currency, trading, online currency trading 12 forex currency trading save 12 currency trading save 12 forex, currency, trading, online currency trading 12 currency trading save 12 currency trading save 12 forex, currency, trading, currency forex trading 12 currency trading save 12 currency trading save 12 currency, trading, currency trading 12 currency, trading, currency trading 12 currency, trading, currency trading 12		currency forex online trading SAVE 12	trading	10
e currency trading save 11 e currency day trading save 12 currency day trading save 12 currency, day, trading, currency day currencytrading, trading 162 forex currency trading system save 11 foreign currency trading save 11 online, currency, trading, system, forex currency trading system 12 foreign, currency, trading, online currency trading 12 online, currency, trading, online currency trading 12 forex currency trading save 12 currency trading save 12 forex currency trading save 12 currency, trading, online currency trading 12 forex, currency, trading, online currency trading 12 currency, trading, currency, trading, currency forex trading 12 forex, currency, trading, currency forex trading 12 currency, trading save 12 forex, currency, trading, currency forex trading 12 currency, trading, currency trading 12 forex, currency, trading, currency forex trading 12 currency, trading, currency trading 12 forex, currency, trading, currency trading 12		currency forex online trading currency online forex trading	recovery,	10
e currency trading save 10 e, currency, trading, e currency trading currency day trading save currency, day, trading, currency day currency trading forex currency trading system save foreign currency trading save foreign currency trading save foreign currency trading save forex currency trading save foreign currency trading save forex currency trading save<		Conteney, Teres, entities, trading, conteney entities (trading	currency,	10
e, currency, trading, e currency trading ▶ All Tags 162 currency day trading sAVE 12 currency, day, trading, currency day currencytrading, trading 11 forex currency trading system sAVE 11 foreign currency trading system sAVE 11 foreign currency trading saVE 12 online currency trading sAVE 12 foreign, currency, trading, foreign currency trading 12 online, currency, trading, online currency trading 12 forex, currency, trading, online currency trading 12 forex, currency, trading, currency trading 12 forex, currency, trading, currency trading 12 currency trading sAVE 12 forex, currency, trading, currency trading 12 forex, currency, trading, currency forex trading 12 currency trading sAVE 12 forex, currency, trading, currency forex trading 12 currency, trading, currency trading 12 forex, currency, trading, currency trading <td></td> <td>e currency trading save</td> <td>debt,</td> <td>10</td>		e currency trading save	debt,	10
currency day trading sAVE 12 currency, day, trading, currency day currencytrading, trading forex currency trading system sAVE 11 foreign currency trading system sAVE 11 foreign currency trading save 12 online currency trading save 12 online, currency, trading, online 12 forex currency trading save 12 online, currency, trading, online 12 forex currency trading save 12 online, currency, trading, online 12 forex currency trading save 12 currency trading save 12 forex, currency, trading, currency forex trading 12 currency trading save 12 forex, currency, trading, currency forex trading 12 currency trading save 12 forex, currency, trading, currency forex trading 12 currency, trading, save 12 forex, currency, trading, currency forex trading 12 currency, trading, currency trading save 12 forex, currency, trading, currency trading 12 forex, currency, trading, currency trading 12 forex, currency, tr		e, currency, trading, e currency trading	All Tags	162
currency, day, trading, currency day currencytrading, trading forex currency trading system SAVE foreign currency trading save foreign currency trading save forex currency trading save online currency trading save online, currency, trading, online forex currency trading save currency trading save online, currency, trading, online forex, currency, trading, currency, trading, online currency trading save forex, currency, trading, currency forex trading forex, currency, trading, currency forex trading currency trading save forex, currency, trading, currency forex trading currency trading save		currency day trading SAVE 12		
forex currency trading system save 11 forex, currency, trading, system, forex forex, currency foreign currency trading save 12 foreign, currency, trading, foreign, currency trading online currency trading save 12 online, currency, trading, online 12 forex currency trading save 12 online, currency, trading, online 12 forex, currency, trading, online 12 forex, currency, trading, save 12 forex, currency, trading, currency forex 12 forex, currency, trading, currency forex 12 forex, currency, trading, currency forex 12		currency, day, trading, currency day currencytrading, trading		
foreign currency trading system save 12 foreign currency trading save 12 online currency trading save 12 online, currency, trading, foreign currency trading forex currency trading save 12 online, currency, trading, online currency trading forex currency trading save 12 forex, currency, trading, currency trading currency trading save 12 forex, currency, trading, currency forex trading currency trading save 12 forex, currency, trading, currency forex trading currency trading save 12 forex, currency, trading, currency forex trading		forev surrange trading system our 11		
torex, currency, trading, system, forex currency trading system foreign currency trading save online currency trading save online, currency, trading, online currency trading forex currency trading save forex currency trading save currency, trading, online currency trading forex, currency, trading, online currency trading forex, currency, trading, currency forex trading currency trading save forex, currency, trading, currency forex trading currency, trading, currency trading save				
foreign currency trading SAVE 12 foreign, currency, trading, foreign, currency trading online currency trading SAVE 12 12 online, currency, trading, online currency forex currency trading SAVE 12 forex, currency, trading, currency currency trading SAVE 12 forex, currency, trading, currency trading SAVE 12		torex, currency, trading, system, torex currency trading system		
foreign, currency, trading, foreign currency trading online currency trading save forex currency trading save forex, currency, trading, currency forex trading currency trading save currency trading save forex, currency, trading, currency forex trading currency trading save currency, trading, currency forex trading currency, trading, currency trading		foreign currency trading SAVE 12		
online currency trading SAVE 12 online, currency, trading, online currency trading forex currency trading SAVE forex, currency, trading, currency forex trading currency trading SAVE currency trading SAVE		foreign, currency, trading, foreign currency trading		
online, currency, trading, online currency trading forex currency trading save 12 currency trading save 12 currency trading save 12 currency trading save 12		online currency frading SAVE		
forex currency trading save 12 currency trading save 12 currency trading save 12 currency trading save 12 currency trading save 12		online currency trading SAVE		
forex currency trading save forex, currency, trading, currency forex trading currency trading save 12		onime, currency, trading, onime currency trading		
forex, currency, trading, currency forex trading currency trading save currency, trading, currency trading, trading, currency trading		forex currency trading SAVE		
currency trading SAVE 12 currency, < trading, < currency < trading		forex, currency, trading, currency forex trading		
currency, < trading, < currency < trading		currency trading SAVE 12		
		currency, trading, currency trading		
4 0 0 0 0 0 4 0 5 0 6 0 7 0 0 0 40 0 Novts		1 2 3 4 5 6 7 9 10 Next>		
		98 Bookmarks		

Spamming-resistant Ranking

- HITS-like algorithm
 - Good and expert users have high quality bookmarks
 - High quality bookmarks are tagged by good users
- Discoverer/Follower Scoring Scheme
 - Discoverers are credited more
 - Followers receive lower score
- Experiments show good results against simulated spammers







- Spamming hinders resource sharing and retrieval on social Web sites
- Spamming on social Web sites comes in different forms, has more targets
- Future anti-spamming techniques need to focus on:
 (1) mixing different strategies
 (2) community-specific method