Advanced Topics
Social Spam, Campaigns, Misinformation and Crowdturfing
(Not part of Exam)
# Large-Scale Social Systems

## Online Social Networking
- Facebook
- Twitter
- LinkedIn

## Social Media
- YouTube
- Flickr
- Digg

## Information sharing communities
- Reddit
- Yahoo! Answers
- StumbleUpon

## Social Games
- Zynga
- Rovio
- EA
- Wooga

## Location-based Services
- Foursquare
- Yelp
- Google Latitude
- Gowalla

## Crowd-based services
- CrowdFlower
- Kickstarter
- IndieGogo
Key Organizing Principles

- **Openness:**
  - Social systems are inherently open to users who generate, share and consume information
  - E.g., post a message, upload and watch a video
- **Collaboration:**
  - Many users organically participate in social systems to engage in collaborative activities
  - E.g., organize for political change, share disaster-related information
- **Real-time information propagation:**
  - Users, media and organization post information related to hot events in (near) real-time
  - E.g., emergency alerts, natural disaster news and sports games
- **Crowdsourcing tasks or hiring cheap workers from all over the world:**
  - People can hire workers from crowdsourcing sites with paying little money
  - E.g., workers from Amazon Mechanical Turk for labeling data, workers from Fiverr for editing a document
Challenges and Research Approach

These necessary positive aspects may also lead to negative consequences

- Spam of many flavors
  - Comment spam (~90% on websites = 46 billion)
  - Spam tweets (1% = 3 million/day) and Twitter spammers (5% = 25 million)
  - Spam videos (20%)  
- Traditional Attacks
  - Phishing, malware and etc
- Campaigns
- Misinformation
- Crowdturfing
- Misuse
  - Crowdsourcing the wrong guy in the Boston bombings at Reddit
- ...

Fake Accounts

9% on Facebook = 87 million accounts in 2012 [Facebook]

[Image of a fake Facebook profile with a note: FAKE]

Send Katherine a Message
Poke Katherine

Information
Birthday: August 29, 1987

Basic Information
Sex: Female
Birthday: August 29, 1987

Personal Information
About Me:
Hi !!!

---------------
http://dokor.com - MY HOME VIDEO
---------------

do not click!
Comment Spam

- **83 ~ 90% on websites = 46 billion comments** [Akismet and Mollom. 2010, Kant et al. WSDM 2012]

---

**Rosiane**
facebook.com/profile.php?id=100003406202721x
m.smeeen@nail.ru
100.143.232.12

Submitted on 2012/07/02 at 09:27
you people may not believe at all but I can and will tell you that between heaven and earth are things beyond the reach of ordinary man and women.you people do not know what knowledge is and you would not gain any knowledge if it's not by some divine revelation.is this the book of the devil maybe but it sure as hell is not for ordinary folks like you people to read, you could not handle it any one of you, before you open the book of the devil you better make sure your in a right pad with GOD Jehova.

---

**Urvi**
facebook.com/profile.php?id=100003406194827x
info@sms-vluchtelingen.nll
188.143.232.12

Submitted on 2012/07/02 at 02:20
I had a spambot at my potrey site post something regarding the size of her husband.All I can say is Mr. Jeremy must be glad he isn’t married to her.Then there’s the one with the guy wanting to sell his bridal dresses.

---

**best affiliate website**
home-businessreviews.com/turnkey-affiliate-websites...
x
justinki111558@gmail.com
46.109.196.107

Submitted on 2012/06/29 at 04:34
Make $1,000’s Weekly with a Health Internet Business of Your Very Own
Now get a complete fully-operational “Health eBiz” in a box!
This amazing site:
* Closes sales automatically for you!
* Has a complete electronic sales manager that makes all up sells for you!
Spam Tweets and Twitter Spammers

1% Spam tweets and 5% Twitter spammers
- 3 million spam tweets/day and 25 million spam accounts

[Twitter and TwitSweeper, 2010]

cphtlink: for all those wondering about the twitter trend i do believe Apple Shampoo refers to a Blink 182 song.
10 minutes ago from web · Reply · View Tweet

TrevBusiness: #apple shampoo Getting strange calls? Reverse Phone Detective to find out who’s bothering you! www.SmallBusinessSolved.com/r/
10 minutes ago from web · Reply · View Tweet

TrevBusiness: #apple shampoo This website to give me great ideas to do on a daily to get the love flowing www.SmallBusinessSolved.com/m/
10 minutes ago from web · Reply · View Tweet
Spam Videos

- 183 million U.S. Internet users watched more than 37 billion online videos in Oct 2012. [comScore]
- 20% of online videos are spam [VideoSurf]
Collective Attention Spam

- Target popular and trendy topics/items
- Feed spam contents once the topics/items become popular

Results for #DrakeCriesWhen

- CherrylKornblut: No way. She pulls this again! bit.ly/Falmo #gritosmexicanos #DrakeCriesWhen #FastFoodAddiction Glen Rice
- AprylHemmelgarn: I wonder if this really works bit.ly/Falmo Glen Rice #UKnowUHungryWhen #gritosmexicanos #DrakeCriesWhen #FastFoodAddiction
- UnSchwegel6358: Anybody know is this really works!? bit.ly/Falmo #gritosmexicanos Glen Rice #UKnowUHungryWhen #DrakocriesWhen #FastFoodAddiction
- DebraRozzi7831: No way. She pulls this again! bit.ly/Falmo #gritosmexicanos #DrakeCriesWhen #FastFoodAddiction Glen Rice
- MaryaLiccketto9: Omg...is this real! bit.ly/Falmo #gritosmexicanos #DrakeCriesWhen #FastFoodAddiction Glen Rice
**Campaigns**

### Astroturfing

The need to protect the internet from 'astroturfing' grows ever more urgent. The tobacco industry does it, the US Air Force clearly wants to... astroturfing - the use of sophisticated software to drown out real people on web forums - is on the rise. How do we stop it?

![A real person using the internet. Unfortunately, we can no longer assume what we are reading is written by one of these creatures. Photograph: Jeff Blackburn/Fax Features](image)

### Fake review campaign

1. **3 of 1 people found the following review helpful:**

   **Practically FREE music**, December 4, 2004
   
   - *This review is from:* Audio Street Pro (CD-ROM)
   - *I can't believe for $10 (after rebate) I got a program that gives me free unlimited music. I was hoping it did half what it was.*

2. **2 of 2 people found the following review helpful:**

   **Like a tape recorder**, December 8, 2004
   
   - *This review is from:* Audio Street Pro (CD-ROM)
   - *This software really rocks. I can set the program to record music all day long and just let it go. I come home and my...*

3. **8 of 8 people found the following review helpful:**

   **Yes - it really works**, December 4, 2004
   
   - *See my review for Audio Street - this PRO is even better. This is the solution I've been looking for. After buying Times,*

4. **3 of 5 people found the following review helpful:**

   **My kids love it**, December 4, 2004
   
   - *This review is from:* Audio Street Pro (CD-ROM)
   - *This was a bargain at $20 - better than the other ones that have me above water costs. My kids get a kick out of the...*

### Political campaign

**Bogus Grass-Roots Politics on Twitter**

Data-mining techniques reveal fake Twitter accounts that give the impression of a vast political movement.

*Source: Wang et al. WWW 2012*

<table>
<thead>
<tr>
<th>Website</th>
<th>Campaigns’ % Crowdfunding</th>
<th>Tasks</th>
<th>$ per Subm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon Turk (US)</td>
<td>41K 12%</td>
<td>2.9M</td>
<td>$0.092</td>
</tr>
<tr>
<td>ShortTask* (US)</td>
<td>30K 95%</td>
<td>527K</td>
<td>$0.096</td>
</tr>
<tr>
<td>MinuteWorkers (US)</td>
<td>710 70%</td>
<td>10K</td>
<td>$0.241</td>
</tr>
<tr>
<td>MyEasy Task (US)</td>
<td>166 83%</td>
<td>4K</td>
<td>$0.149</td>
</tr>
<tr>
<td>Microworkers (US)</td>
<td>267 89%</td>
<td>84K</td>
<td>$0.175</td>
</tr>
</tbody>
</table>

Researchers have found evidence that political campaigns and special-interest groups are using scores of fake Twitter accounts to create the impression of broad grass-roots political expression. A team at Indiana University used data-mining and network-analysis techniques to detect the activity.

"We think this technique must be common," says Filippo Menczer, an associate professor at Indiana University and one of the principal investigators on the project. "Wherever there are lots of eyes looking at screens, spammers will be there; so why not with politics?"
Adversarial Propaganda

- Create and spread rumors and Misinformation
- Target a product/ government

Pentagon Wants a Social Media Propaganda Machine

BY ADAM RAWNSLEY 07.15.11 2:40 PM

You don’t need to have 5,000 friends of Facebook to know that social media can have a notorious mix of rumor, gossip and just plain disinformation. The Pentagon is looking to build a tool to sniff out social media propaganda campaigns and spit some counter-spin right back at it.

On Thursday, Defense Department extreme technology arm Darpa unveiled its Social Media in...
Misinformation (Fake)

Fake Images
Crowdturfing (Crowdsourcing + Astroturfing)

- A multimillion-dollar industry in Chinese crowdsourcing sites
  - 90% crowdturfing tasks [MIT Technology Review]
- 70~95% crowdturfing tasks at several U.S. crowdsourcing sites [Wang et al., WWW 2012]

**Twitter Post: CPP Scam**

Work done: 222/250
You will earn $0.60
This task takes less than 30 min to finish
Job ID: 364488d297e8

What is expected from Workers?

You must have 50 Twitter followers. Make sure you are logged into your Twitter account

1. Open your browser and search on Google "college pro painters success"
2. Click on any search result that starts with collegepropaintersscam.com
3. Go to Home Page of the website
4. Retweet any article
Examples of Crowdturfing

- Vietnamese propaganda spread by 1,000 crowdturfers

Vietnam admits deploying bloggers to support government

By Nga Pham
BBC News, Hanoi

Vietnamese propaganda officials have admitted deploying people to engage in online discussions and post comments supporting the Communist Party's policies.

The party has also confirmed that it operates a network of nearly 1,000 "public opinion shapers".

They are assigned with the task of spreading the party line.

The tactic is similar to China's model of internet moderators who aim to control news and manipulate opinion.

"Political opportunists"

Hanoi Propaganda and Education Department head Ho Quang Loi said that the authorities had hired hundreds of so-called "internet polemists" in the fight against "online hostile forces".

The bloggers have been hailed for stopping negative online rumours
Social Spam

- **Fake accounts (5 ~ 6% on Facebook = 42 million)**
  - [Facebook. 2012]
- **Comment spam (83 ~ 90% on websites = 46 billion)**
  - [Akismet and Mollom. 2010, Kant et al. WSDM 2012]
- **Spam Tweets (1% = 3 million/day) and Twitter Spammers (5% = 25 million)**
  - [Twitter. 2010, TwitSweeper. 2010, Lee et al. SIGIR 2010, Lee et al. ICWSM 2011, Yang et al. WWW 2012]
- **Tag spam**
- **Spam videos**
  - [Benevenuto et al. AIRWeb 2008, Benevenuto et al. SIGIR 2009]
- **Fake Reviews**
  - [Jindal and Bing ICDM 2007, Lim et al. CIKM 2010, Wang et al. TIST 2011, Mukherjee et al. WWW 2012]
- **Voting spam**
  - [Bian et al. AIRWEB 2008, Tran et al. NSDI 2009]
- **Wikipedia vandalism**
- ...
How to detect these?
Blacklisting URLs

- Crawled URLs from Twitter
  - 25 million URLs crawled
  - 8% of them link to spam pages

- Over 80% of spam URLs were shortened
  - Mask landing site
  - Defeat blacklist filtering
    - bit.ly -> short.to -> malware landing page

Grier, C., Thomas, K., Paxson, V., and Zhang, M. @spam: the underground on 140 characters or less. In CCS, 2010.
Blacklist Performance

- Blacklists are slow to list spam domains
  - 80% of clicks are seen in first day

- Retroactively blacklist

![URIBL and Google malware graphs showing the number of tweets over days with red and blue colors indicating lead and lag.](chart.png)
Comparison to Email Click Through

- Spam Email clickthrough: .003-.006%
  - From Spamalytics, Kanich et al. CCS 2008

- Twitter clickthrough: .13%
  - Collected 245,000 spam URLs
  - Define clickthrough as clicks / reach
  - Reach defined as tweets * followers
Social Spam Detection Approaches

• Supervised spam detection approach
  – The most popular approach
  – Require labeled data for training purpose

• Ranking users based on their social graph

• Use crowd wisdom (humans) to identify fake accounts
Open Research Challenges

• Need for large, accurate, up-to-date data sets
  – APIs
  – Hard crawling
  – Shared datasets
  – Purchasing data (e.g., Gnip)
  – Data grant or know an insider

• Labeling
  – Manual labeling
  – Use crowd wisdom
  – Get labeled data from a social media site
  – Blacklist
Open Research Challenges

• Integration of multiple techniques for data processing and modeling
  – Big data analysis, machine learning (data mining), information retrieval, visualization, etc

• Interdisciplinary research for analysis
  – computer science, social science, psychology, etc

• Arms race (endless battle)
  – Spammers and malicious users change their behaviors or use new techniques to avoid existing detection approaches
  – Spammers and malicious users move to another site
Useful Tools

• Machine learning
  – LingPipe (linguistic analysis): http://alias-i.com/lingpipe/

• Visualization
  – Matplotlib: http://matplotlib.org/
  – Gephi: https://gephi.org/
  – Graphviz: http://www.graphviz.org/
Useful Tools

• Big data analysis and visualization
  – Hadoop (MapReduce): http://hadoop.apache.org/
  – Pig: https://pig.apache.org/
  – Hive: https://hive.apache.org/
  – Cascalog: http://cascalog.org/
  – Giraph: https://giraph.apache.org/

• Scalable machine learning
  – Mahout: https://mahout.apache.org/

• Large scale stream processing
  – Storm: http://storm.incubator.apache.org/
  – Summingbird: https://github.com/twitter/summingbird
Web 3.0
Web 1.0

Web 1.0
Media companies put content in the web and push it to users.

Low content variety
Content production is slow and expensive

Professional authors
Create content based on what they think people want

Push!

Managers
Hand-manage publishing businesses

Users = consumers
Users consume professionally produced content
Web 1.0

web1.0 problem
This old media model ignored the new possibilities of the internet.

Problems

10,000 editors serve 500,000,000 internet users

Limited content

Limited creativity

Limited business
Web 2.0

web2.0
New platforms allow users to generate content themselves

User generated content
Blogs, videos, photos, music etc.

High content variety
dramatically increases, while technical quality goes down.

Authoring platforms
enable everyone to publish!

Web2.0 managers
are hand-managing the platforms

Professional authors
can still publish, but are competing with – everyone!

Users = authors
users generate content

Yes, we can!
Web 2.0

YOU CAN PUBLISH!
Web 2.0

User generated content

We all serve each other

unlimited content
Basically, web 2.0 is a social change. The technical part of the web hasn’t changed very much.
But, web 3.0 will be driven by technological changes.
Web 3.0 - the semantic web - is about the meaning of data.
Suppose, I am a stamp collector...
Over the years I’ve collected a lot of stamps.
About every stamp, I made a document
Web 3.0

That’s a lot of documents
Web 3.0

How will I find that specific stamp?
Ah, of course!
This is the web we have today: a huge collection of documents
The words of all those documents are indexed. We can search for keywords.
Web 3.0

Now, suppose I search for all red stamps...
Web 3.0

What do I get?

Red stamps
Stamps from Cambodia (Khmer Rouge)
Stamps from the Red Sea
Stamps from the 140th anniversary of the Red Cross
Stamps with red dragons
Not very intelligent, but how can a computer know what I mean?
Answer: when we structurally describe that a stamp is a stamp and red is a color.
Describing data in a structured way can best be done in a database.
A database with stamps
A database with countries
A database with colours
A database with stamp traders
Web 3.0 creates a big collection of databases which can be connected on demand.
Agreements are made on the structure of data and the way data is described. Where the data is located is irrelevant. Linking data is the power of web 3.0.
So, “I want all the red stamps, designed in Europe, but used in the U.S.A., between 1980 and 1990” is a question that will get a better answer with web 3.0.
This definition of web 3.0 is a ‘narrow’ definition. Like web 2.0, web 3.0 stands for a range of developments.
A broader definition of web 3.0:

A fast broadband connection to the internet, always and everywhere.

Open source techniques and free data (Data as a Service)

Open identities

Software as a Service (e.g. Google docs)
References