ONLINE SOCIAL NETWORK ANALYSIS

Emrah Cem and Kamil Sarac
University of Texas at Dallas

Copyright -- Emrah Cem and Kamil Sarac
OUTLINE

- What is an Online Social Network (OSN)?
- When did they emerge?
- Privacy issues in OSNs
- Why is it important to analyze?
- What are the challenges?
- Demo
- Take aways…
What is Online Social Network (OSN)?

- When a computer network connects people or organizations, it is called OSN.
Skype Statistics (http://www.statisticbrain.com/skype-
statistics/)
Total number of Skype users as of Jan. 2012
31 million
Total percentage of small businesses that use Skype as
primary communication service
35%
Number of iPhone Skype downloads in 2010
7 million
Skype revenue in 2010
$406.2 million
Number of monthly log-ins to Skype
124 million
Amount of money spent by Microsoft to acquire Skype
$8.5 billion

Statistics)
Number of Wikipedia Articles
21.2 million
Global Unique Visitors
482 million
Page requests for February
18.1 billion
New articles per day
8.449
Wikipedia Edits Per Month
12.4 million
New Editors Per Month for All Wikimedia Projects
16,170

Facebook Statistics (http://www.searchenginejournal.com/stats-
on-facebook-2012-infographic/40301/)
Number of active users
~1 billion
Daily active users
584 million
Total number of connections
~100 billion
Number of photos uploaded daily
250 million
Number of mobile users
425 million

Twitter Statistics (http://www.website-monitoring.com/blog/2012/11/07/twitter-
2012-facts-and-figures-infographic/)
Number of registered users
500 million
Number of active users
140 million
New visitors everyday
300,000
Number of tweets per second
750
Number of daily mobile users
6.9 million

LinkedIn Statistics (http://press.linkedin.com/about)
Number of active users
~190 million
New members per second
2
Number of searches in 2012
5 billion
New visitors everyday
300,000
Number of tweets per second
750
Number of daily mobile users
6.9 million
New Editors Per Month for All Wikimedia Projects
16,170
When did it emerge?

1997 SixDegrees.com
1998
1999 Live Journal Asia Avenue Black Planet
2000 Lunarstorm MiGente SixDegrees - closes!
2001 Cyworld Ryze
2002 Fotolog Friendster Skyblog
2003 Couchsurfing LinkedIn MySpace Tribe.net Last.FM Hi5
2004 Orkut, Dogster Flickr, Piczo, Mixi, Facebook (Harvard) Multiply, a SmallWorld Dogeball, Care2 Catster Hyves
2005 Yahoo 360 YouTube, Xanga Cyworld Bebo Facebook (High School) Ning Asian Ave
2006 Facebook (work) Windows Live Cyworld (US) Twitter MyChurch, Facebook (Everyone)

Copyright -- Emrah Cem and Kamil Sarac
Why do “we” generally use OSNs?

- To keep in touch with our friends
- To make new contacts
- To find people with similar interests and ideas

Figure 1 The Facebook Iceberg Model (Iceberg image © Ralph A. Clevenger/CORBIS)
Privacy Issues in OSNs

- Is the personal data only used for marketing and advertising? **No!**
- Identity theft (e.g., via social engineering)
- Use of personal data by third-parties
  - Have you ever seen advertisements of products you have searched on Amazon before while browsing another web page?
- Even stalking (if you share location information)
Is OSNs harmful at all?

- You can find a job (e.g., LinkedIn)
- You can talk to friends using VoIP cheaply (e.g., Skype)
- You can find your old friends (e.g., Facebook)
- Government can find out criminals through OSNs
  - A criminal, accused of bank fraud, started posting on Facebook. Although his profile was private, his list of friends was not, and through this vector, they eventually caught him.
YOU NEED TO TAKE SOME ACTIONS FOR PRIVACY
Is private really “private”?

- Sometimes private information can be inferred using partially available data, so actually it may not always be the case that what you set as private is actually private.
- OSN analysis techniques can be used to figure out some private information.
What does OSN analysis focus on?

- Finding patterns of relations among people, organizations, states, etc.

- Generation Models
- Communities
- Better Infrastructure (for service providers)
- Better recommendations (for service providers)
- Better advertisements (for online commerce)
- Better search results (for search engines)
- Nation’s feelings about president candidates!

Copyright -- Emrah Cem and Kamil Sarac
What is inside OSN Analysis?

**Static Analysis**
- Local Characteristics
  - Degree Distribution
  - Clustering Coefficient
  - Eccentricity
  - Closeness Centrality
  - Betweenness Centrality
  - Pagerank
- Global Characteristics
  - Avg. Degree
  - Avg. Shortest Path
  - Avg. Clustering Coefficient
  - Density
  - Diameter
  - ...
- Community Detection
  - ...

**Dynamic Analysis**
- How do network characteristics change over time?
- How does information disseminate?
- Link/Attribute Prediction
- Tracking Communities
  - ...

Copyright -- Emrah Cem and Kamil Sarac
Take aways

- Take some actions for your privacy
  - [http://blog.eset.com/2012/05/15/millions-have-not-reviewed-facebook-privacy-settings-heres-how](http://blog.eset.com/2012/05/15/millions-have-not-reviewed-facebook-privacy-settings-heres-how)
- OSNs are useful if you worry about your privacy, otherwise it is risky.
- You can do OSN analysis to find out patterns that are not easy to see.
References


- http://blog.eset.com/2012/05/15/millions-have-not-reviewed-facebook-privacy-settings-heres-how