this lecture

introduction & definitions

a human-centric review of research on facebook
  - descriptive analysis of users
  - user motivations
  - user identity
  - social interaction among users
  - the real-name web: privacy & information disclosure
### facebook basic statistics (as of Dec 31, 2019)

http://newsroom.fb.com/company-info/

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mobile month active users</td>
<td>945 M</td>
<td>1.19 B</td>
<td>1.44 B</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>daily active users</td>
<td>757 M</td>
<td>890 M</td>
<td>1.04 B</td>
<td>1.4 B</td>
<td>1.52 B</td>
<td>N/A</td>
</tr>
<tr>
<td>daily active users out US</td>
<td>81%</td>
<td>82.4%</td>
<td>83.6%</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>employees</td>
<td>6337</td>
<td>9199</td>
<td>12691</td>
<td>25105</td>
<td>35587</td>
<td>44942</td>
</tr>
</tbody>
</table>

### definitions: social network site

“web-based [and mobile] services that allow individuals to

1. construct a public or semi-public profile within a bounded system

2. articulate and make visible a list of other users with whom they shared a connection

3. view and traverse their connections and those made by others within the system

the nature of these connections may vary from site to site… meeting strangers is possible but not necessarily the primary practice”

d. boyd and N. B. Ellison, “Social Network Sites: Definition, History and Scholarship,”
understanding research on facebook
(Wilson et al, 2012)

A Review of Facebook Research in the Social Sciences

Robert E. Wilson¹, Samuel D. Gosling², and Lindsay T. Graham²
Department of Psychology, Washington University & Los Alamos National Laboratory, Los Alamos, NM

Abstract

With over 800 million active users, Facebook is changing the way hundreds of millions of people relate to one another and access information. A rapidly growing body of research has accompanied the network growth of Facebook as social psychologists assess the impact of Facebook on social life. In addition, researchers have recognized the utility of Facebook as a novel tool to examine hypotheses in social-scientific, test-scientific, and natural-language processes. However, when one looks at the research literature from a wide variety of disciplines, one finds that it is difficult to keep track of the various papers. And because Facebook is a relatively recent phenomenon, scientists still debate about the best effective ways to do Facebook research. To address these issues, the authors conducted a comprehensive literature search. Identifying 141 relevant articles, which were sorted into 2 categories: descriptive analysis of users, motivations for using Facebook, identity presentation, roles of Facebook in social interactions, and privacy and information disclosure. The literature review serves as the foundation from which to assess current trends and offer recommendations for future research on Facebook and social networking more broadly.

412 articles (end 2011)

"...the questions, methods, and perspectives were so diverse and fragmented that it would be impossible to write a coherent summary of the literature."

"...but we also realized that without summarizing the current trends the situation was unlikely to improve."

research areas on facebook

<table>
<thead>
<tr>
<th>Area of research</th>
<th>No. of articles</th>
<th>% of total</th>
<th>Associated research question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive analyses of users</td>
<td>97</td>
<td>24%</td>
<td>Who is using Facebook, and what are users doing while on Facebook?</td>
</tr>
<tr>
<td>Motivations for using Facebook</td>
<td>78</td>
<td>19%</td>
<td>Why do people use Facebook?</td>
</tr>
<tr>
<td>Identity presentation</td>
<td>50</td>
<td>12%</td>
<td>How are people presenting themselves on Facebook?</td>
</tr>
<tr>
<td>Role of Facebook in social interactions</td>
<td>112</td>
<td>27%</td>
<td>How is Facebook affecting relationships among groups and individuals?</td>
</tr>
<tr>
<td>Privacy and information disclosure</td>
<td>75</td>
<td>18%</td>
<td>Why are people disclosing personal information on Facebook despite potential risks?</td>
</tr>
<tr>
<td>Total</td>
<td>412</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>
Facebook users & articles: growth over time

Fig. 1. Facebook users and articles: Cumulative totals by year

Topic 1: Descriptive analysis of users
Who uses Facebook and what do they do?
graphs as models of networks: quick basics

All materials of this section are taken from:

(a) A graph on 4 nodes.

(b) A directed graph on 4 nodes.

Figure 2.1: Two graphs: (a) an undirected graph, and (b) a directed graph.
**paths and cycles**

**Path:** sequence of nodes where each consecutive pair is connected by an edge

**Simple path:** paths that contain no repeated nodes

**Cycle:** a path with at least three edges, in which the first and last nodes are the same, but otherwise all nodes are distinct
**connectivity**

**Connected component**: a subset of the nodes such that:
(i) every node in the subset has a path to every other;
(ii) the subset is not part of some larger set with the property that every node can reach every other.

(i) says that the component is internally connected
(ii) says that the component is a free-standing piece of the graph

**giant components**

**Giant component**: connected component that contains a significant fraction of all the nodes in large networks

When a network contains a giant component, it almost always contains only one
**path length and distance between nodes**

**Path length**: number of edges in the sequence that comprises it

**Distance between two nodes**: length of the shortest path between them

\[
l(MIT, BBN, RAND, UCLA) = 3 \\
l(MIT, UTAH) = 1 \\
d(LINC, SRI) = 3
\]

**breadth-first search: a method to determine distances**

It searches a graph from a starting node, reaching the closest nodes first.

It serves as a conceptual framework to organize a graph's structure, arranging the nodes based on their distances from a fixed starting point.
small-world phenomenon

**six degrees of separation**
- any two people are separated by no more than six intermediate connections
- the world looks “small” given these short paths
- proposed by Frigyes Karinthy in short story (1929)
- popularized by John Guare’s *Six Degrees of Separation* play (1990)

“I read somewhere that everybody on this planet is separated by only six other people. Six degrees of separation between us and everyone else on this planet.”

**Stanley Milgram’s small-world experiment (1960s)**

- 296 volunteers living in US midwest
- they sent a message to a person living in Boston suburbs
- volunteers could not send message directly to target (unless personal contact)
- they sent message to a personal contact who was likely to know the target
- 64 chains reached the target
- avg. # of intermediate persons = 5.2
  (distance: 6.2)

people navigate short paths with success, even if entire network is not visible
the anatomy of facebook
(Backstrom et al. 2011)

721 million facebook users
69 billion friendships
largest social network ever studied

research questions:
1. how many friends do people have?
2. how many intermediaries between any two users?

L. Backstrom, Anatomy of Facebook, 21.11.2011


J. Ugander, B. Karrer, L. Backstrom, C. Marlow, The Anatomy of the Facebook Social Graph, Nov. 2011,
http://arxiv.org/abs/1111.4503

how many friends?
(Backstrom et al. 2011)

cumulative degree distribution: % of people with less than X friends
10% of people: less than 10 friends
20% of people: less than 25 friends
50% (the median): less than 100 friends
average friend count = 190
4 degrees of separation
(Backstrom et al. 2011) (degrees of separation = intermediaries = distance-1)

average distance (2008) = 5.28
average distance (2011) = 4.74

within same country, connectivity is even higher
84% of all connections are between users in the same country

4 degrees of separation, revisited 2016
https://research.facebook.com/blog/three-and-a-half-degrees-of-separation/

average distance (2008) = 5.28
average distance (2011) = 4.74
average distance (2016) = 4.57

recomputed on entire
FB graph: 1.59 B users
**homophily in facebook**  
(Backstrom et al 2011)

users’ friends are most likely to be of similar age

**homophily**
“tendency of individuals to associate and bond with similar others” (wikipedia)

*birds of a feather flock together*

---

**small-world phenomenon in MS instant messenger**  
(Leskovec et al 2008)

*Figure 2.11: The distribution of distances in the graph of all active Microsoft Instant Messenger user accounts, with an edge joining two users if they communicated at least once during a month-long observation period [27].*

questions?

daniel.gatica-perez@epfl.ch