Introduction to Graph ML

Graph ML

Department of Computer Science
University of Massachusetts, Lowell

Hadi Amiri
hadi@cs.uml.edu

Instructor

- Hadi Amiri
  - DAN-334
  - Office hours
    - by appointment
  - Hobbies: Sports that are hard on the feet!

hadi@cs.uml.edu
www.cs.uml.edu/~hadi

What's This Course about?

- Graph Machine Learning
  - Networks
    - a pattern of inter-connections among a set of things!
    - deal with structure
  - Meta data
    - deal with various user generated content (text, images, videos, etc.,) in networks.

- We aim to learn about prediction algorithms that work well on networks.
  - Models, properties, design principles!

Graphs/Networks

- Communication Networks
  - Telco Nets
  - Messenger Nets
- Friendship Networks
  - Facebook
- Microblogs
  - Twitter
- Information Networks
  - Web!

Examples

Sample 1.

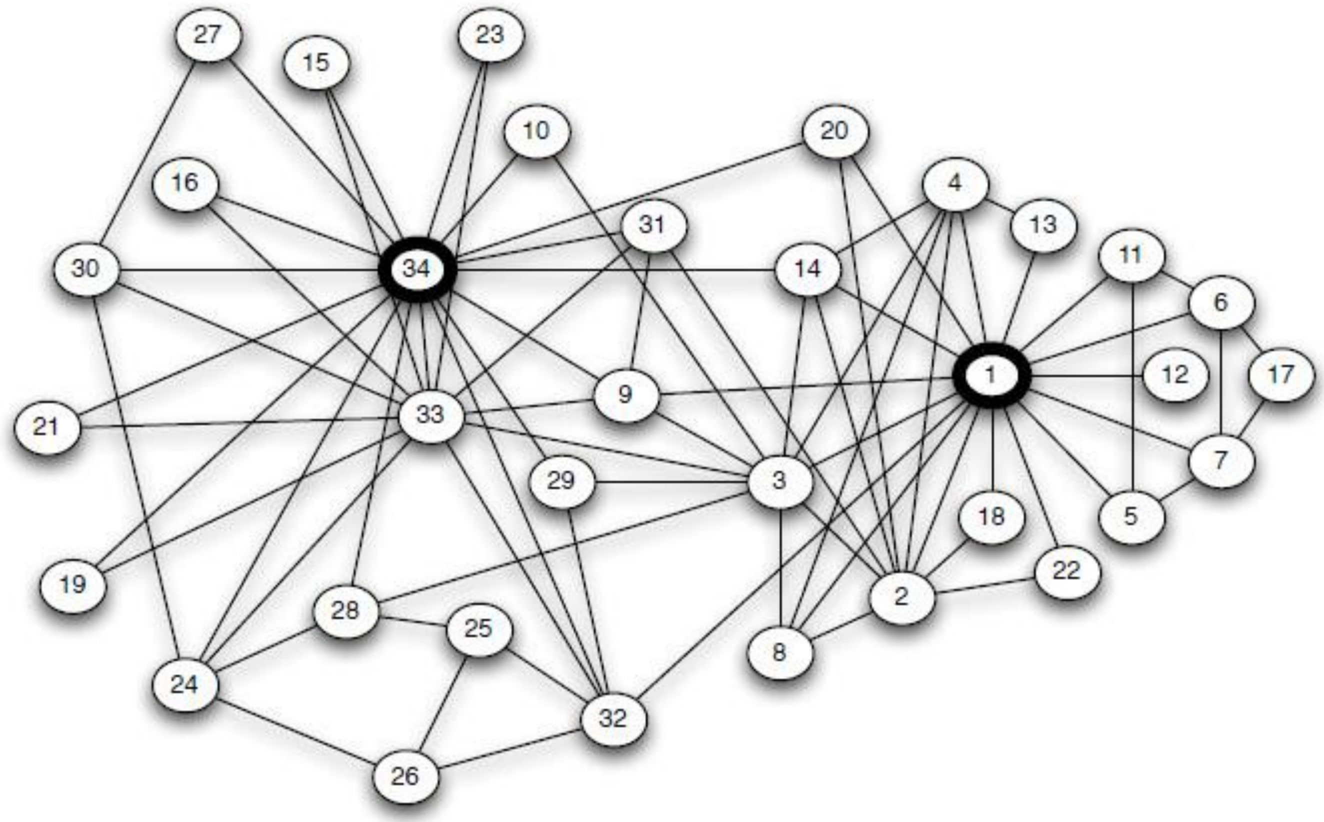


Figure 1.1: The social network of friendships within a 34-person karate club [421].

Sample 2.

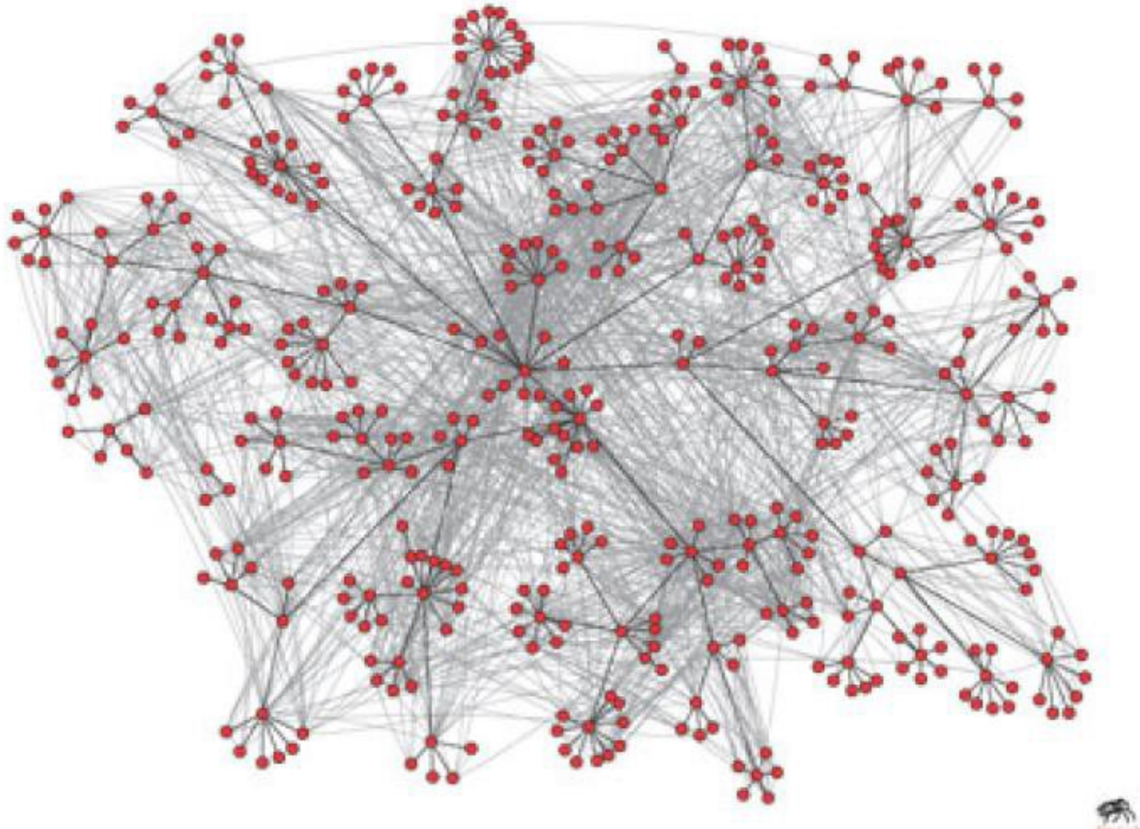


Figure 1.2: Social networks based on communication and interaction can also be constructed from the traces left by on-line data. In this case, the pattern of e-mail communication among 436 employees of Hewlett Packard Research Lab is superimposed on the official organizational hierarchy [6]. (Image from http://www-personal.umich.edu/ ladamic/img/hplabsemailhierarchy.jpg)

Sample 3.

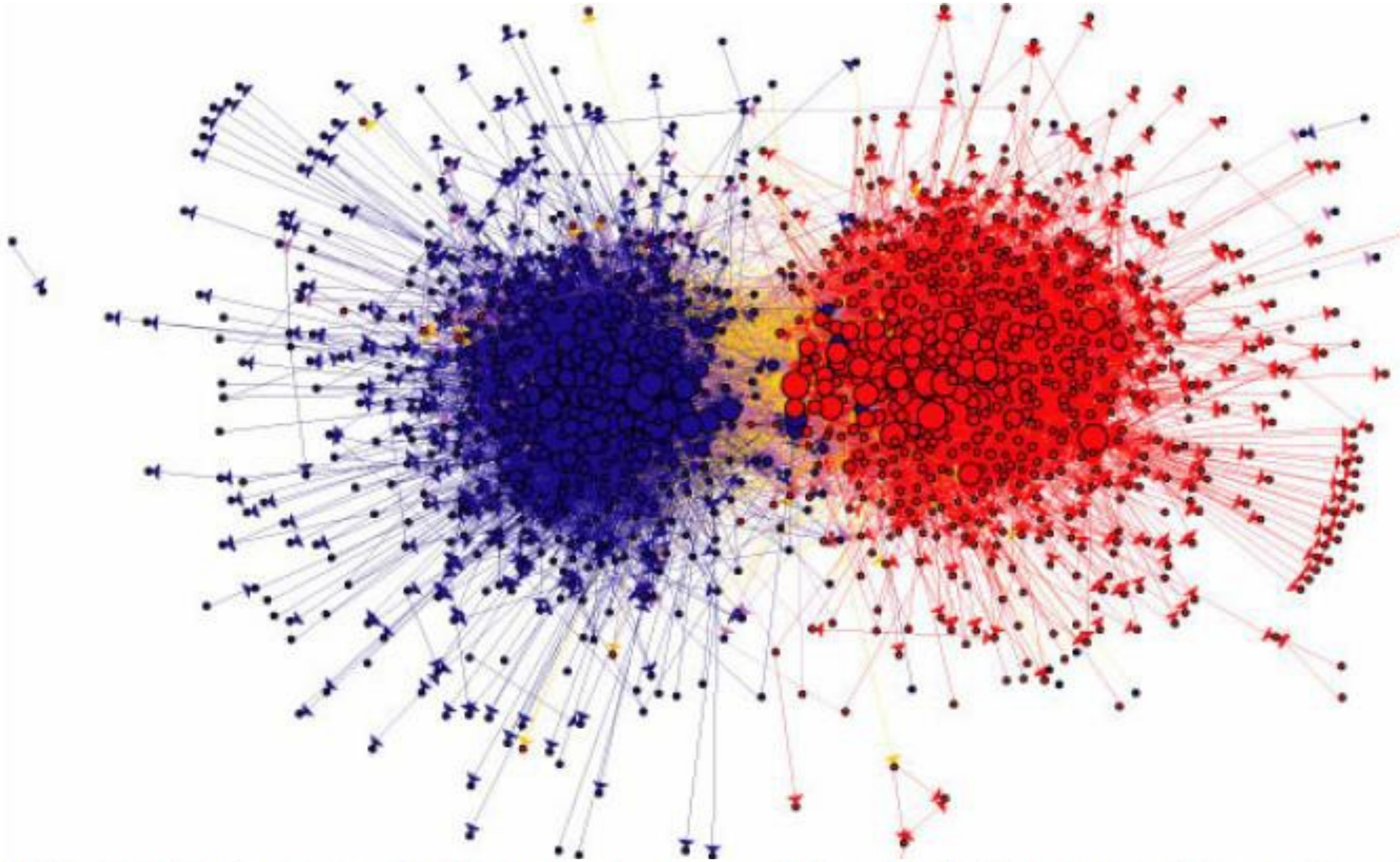


Figure 1.4: The links among Web pages can reveal densely-knit communities and prominent sites. In this case, the network structure of political blogs prior to the 2004 U.S. Presidential election reveals two natural and well-separated clusters [5]. (Image from http://www-personal.umich.edu/ ladamic/img/politicalblogs.jpg)

Sample 4.

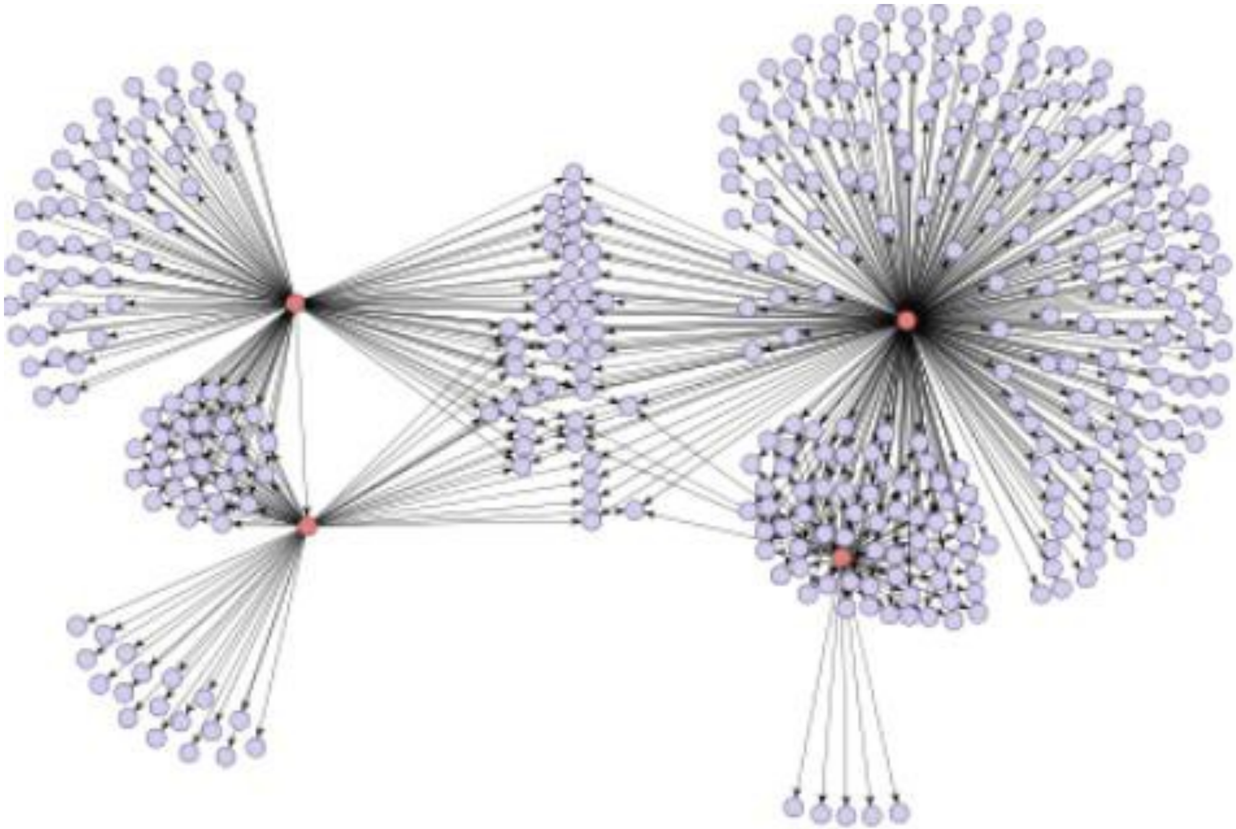


Figure 1.11: When people are influenced by the behaviors their neighbors in the network, the adoption of a new product or innovation can cascade through the network structure. Here, e-mail recommendations for a Japanese graphic novel spread in a kind of informational or social contagion. (Image from Leskovec et al. [271].)

Sample 5.

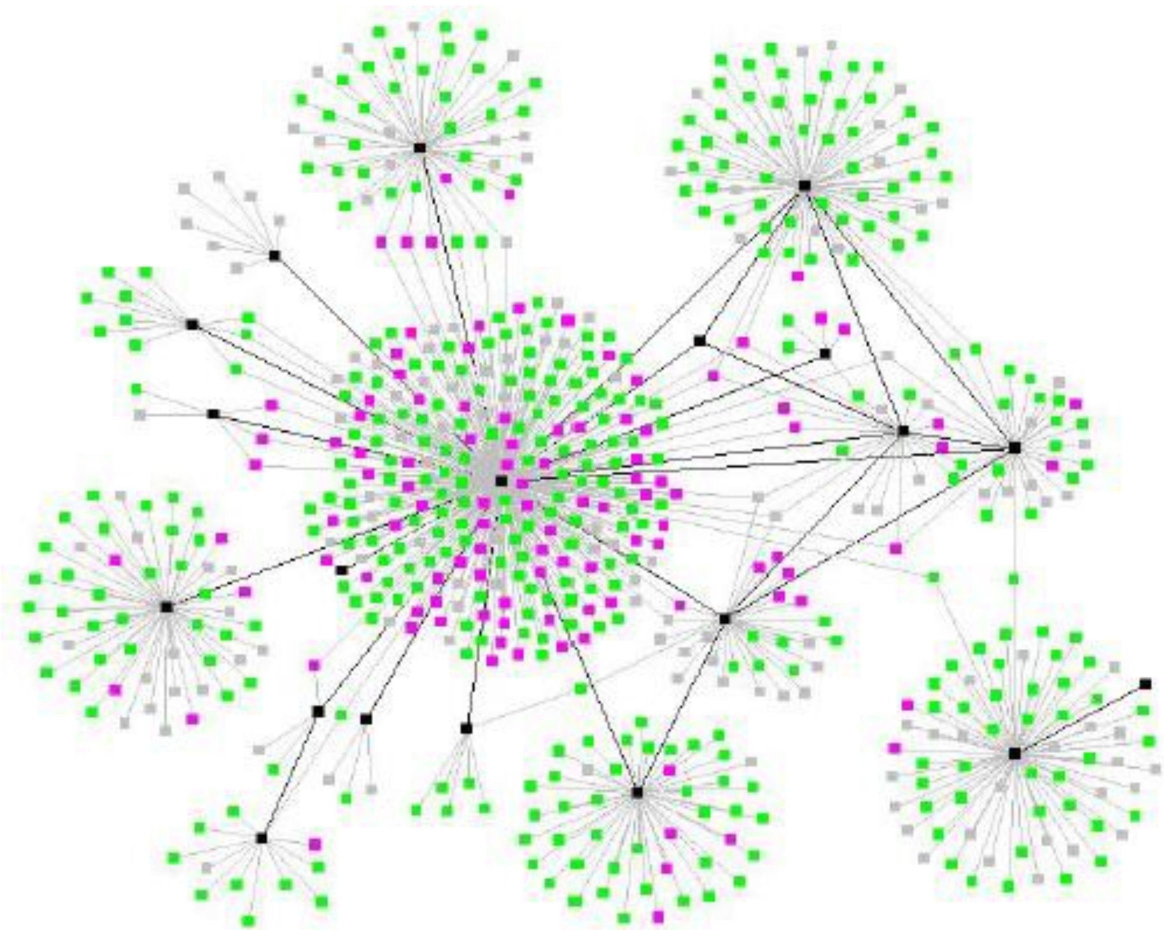
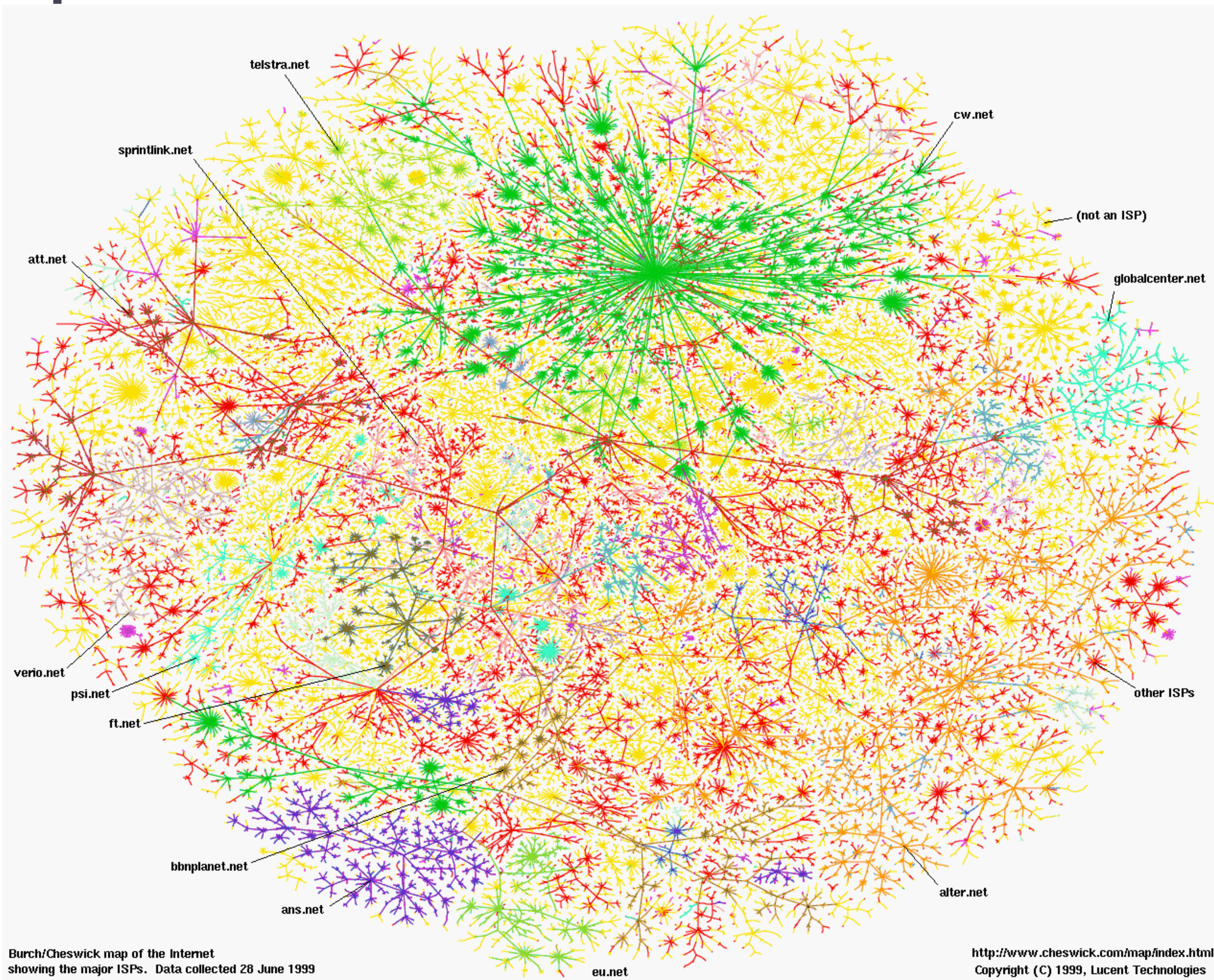


Figure 1.12: The spread of an epidemic disease (such as the tuberculosis outbreak shown here) is another form of cascading behavior in a network. The similarities and contrasts between biological and social contagion lead to interesting research questions. (Image from Andre et al. [16].)

Sample 6.

Network of Major ISPs.
1999



Sample 7.

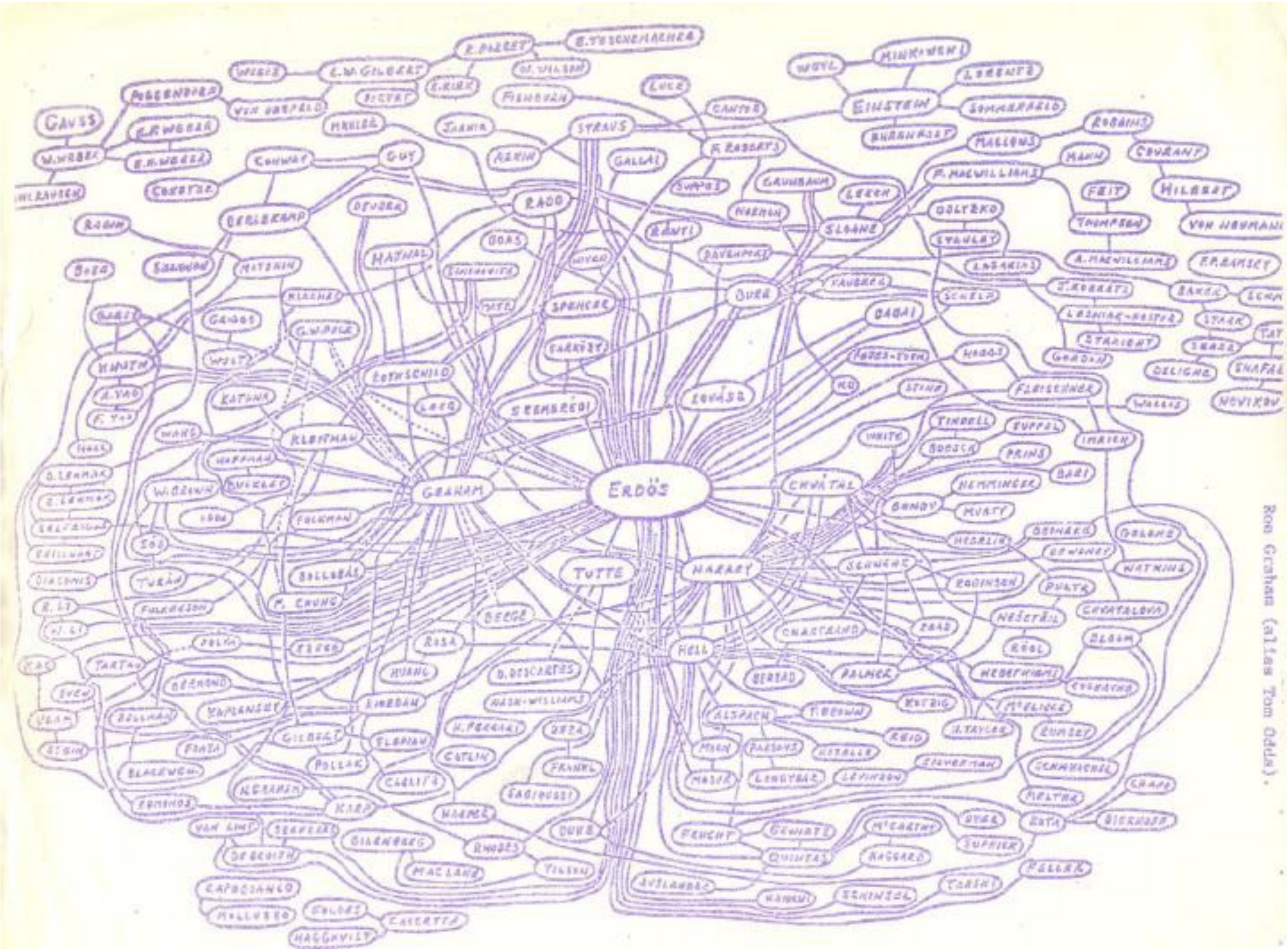


Figure 2.12: Ron Graham's hand-drawn picture of a part of the mathematics collaboration graph, centered on Paul Erdös [189]. (Image from http://www.oakland.edu/enp/cgraph.jpg)

Sample 8.

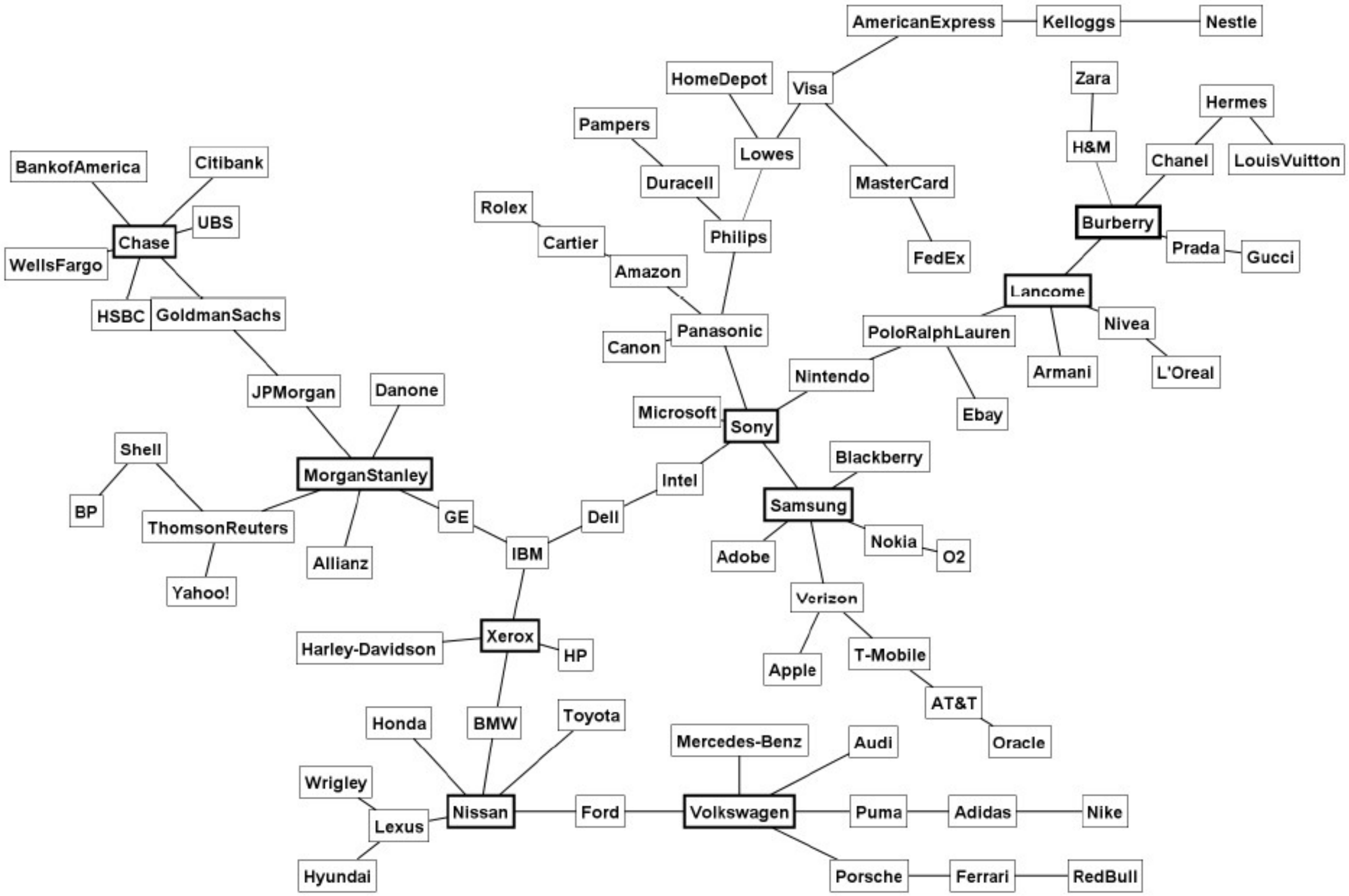


Figure 3. Minimum spanning tree (MST) of the most valued global brands. The MST of the brand network is the subset of edges that forms a tree reaching every brand such that the total length of all the edges is minimized. It is readily apparent that certain brands stand out prominently as hubs with connections to other brands radiating out from them. These hubs are generally the centers of well-formed market category groupings.

Sample 9.

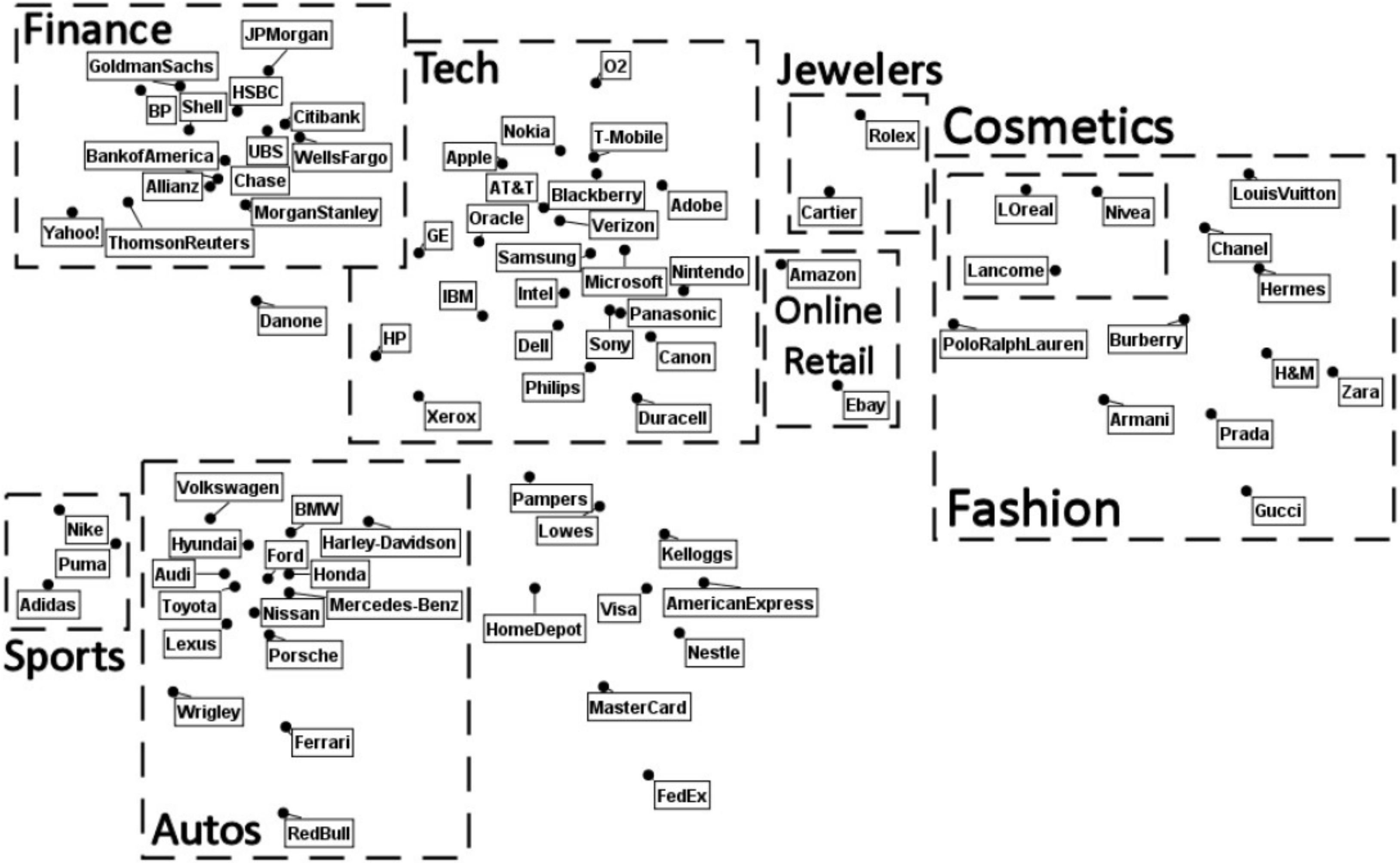
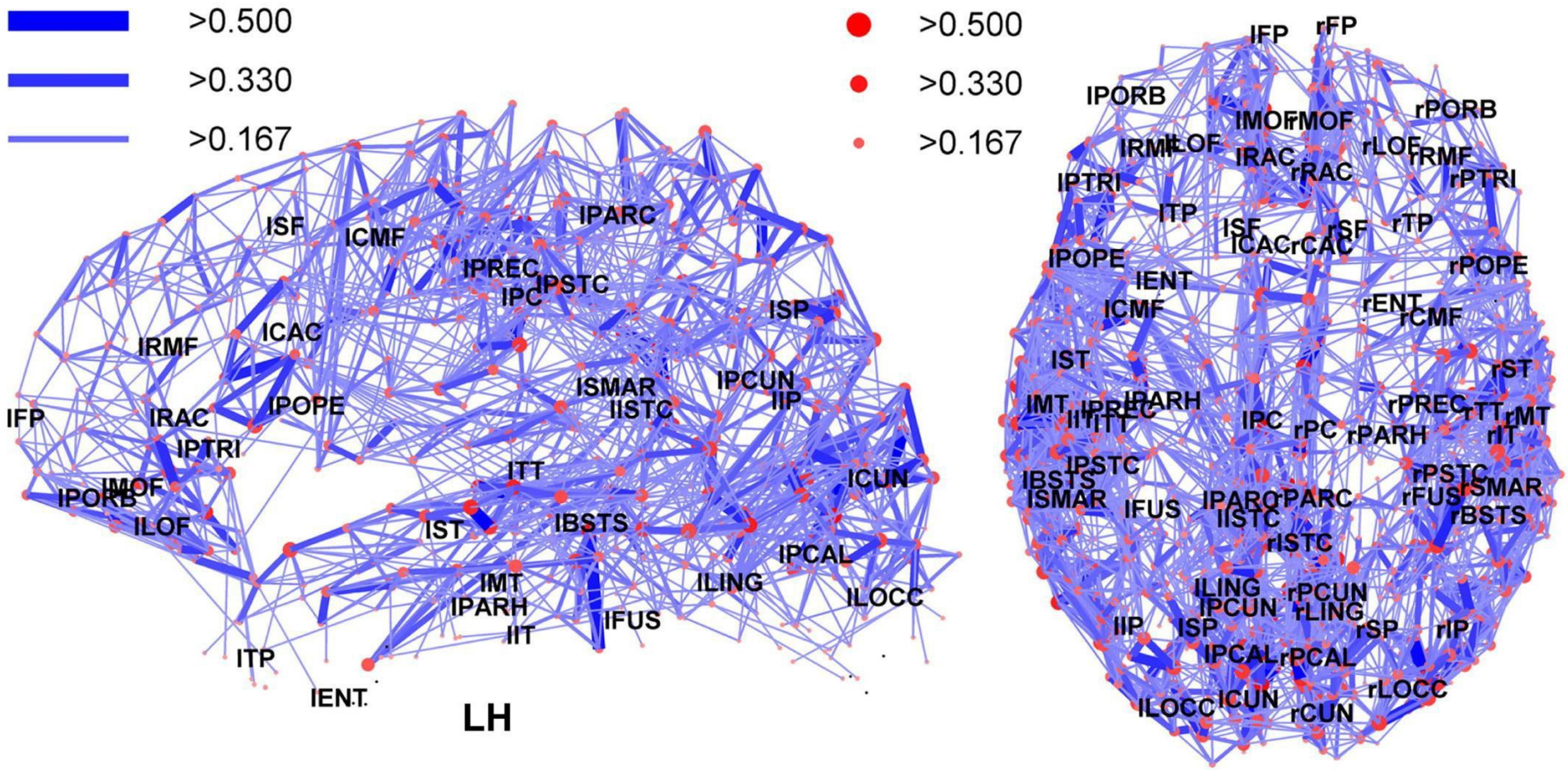


Figure 4. Map of brands. The minimum spanning tree augmented by triangulating each brand location from their nearest neighbors with forced-based layout yields a map high in face validity. Note the eight strong market category groupings outlined with broken lines.

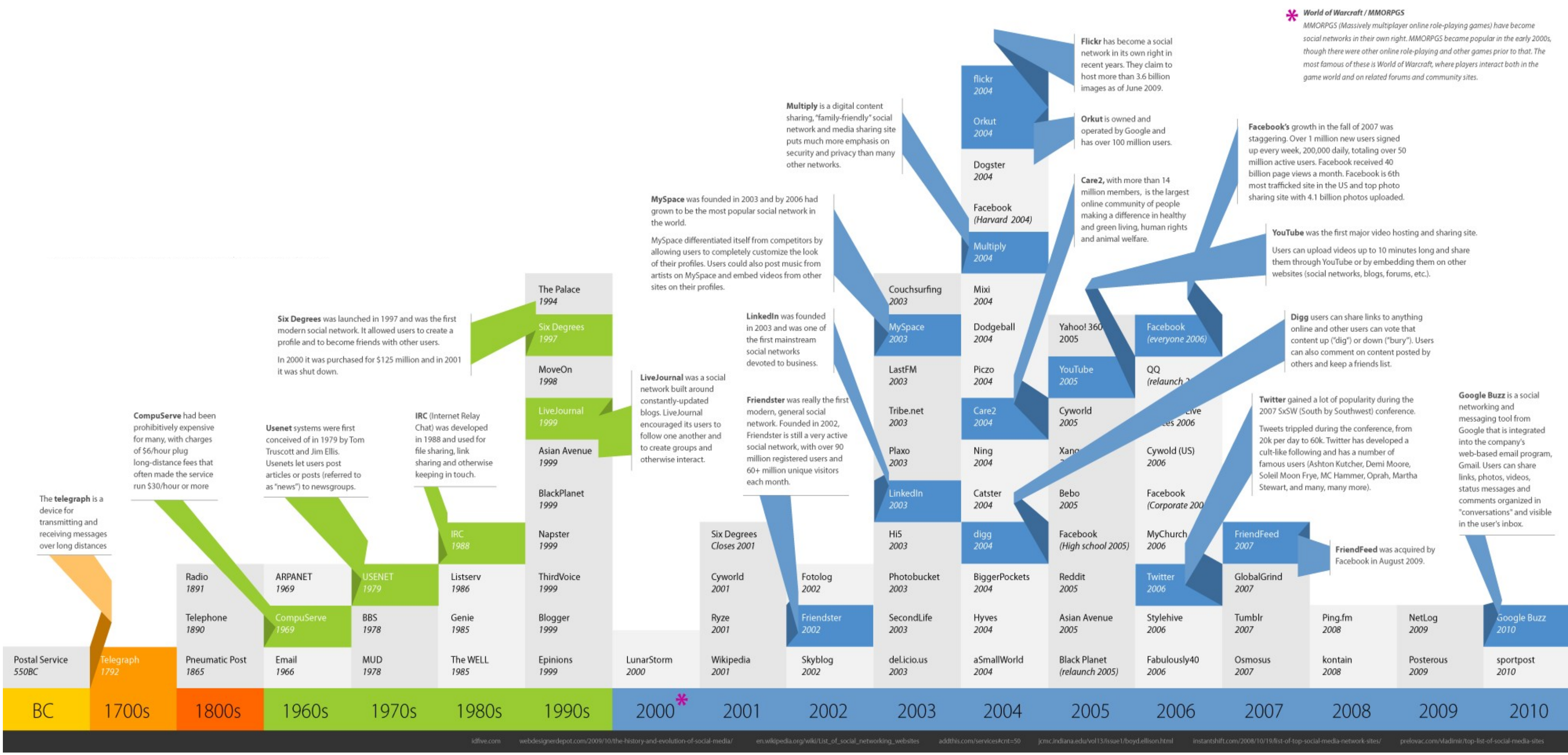
Source: Inferring brand proximities from user-generated content. Paul Dwyer. J of Brand Management 19, 2012.

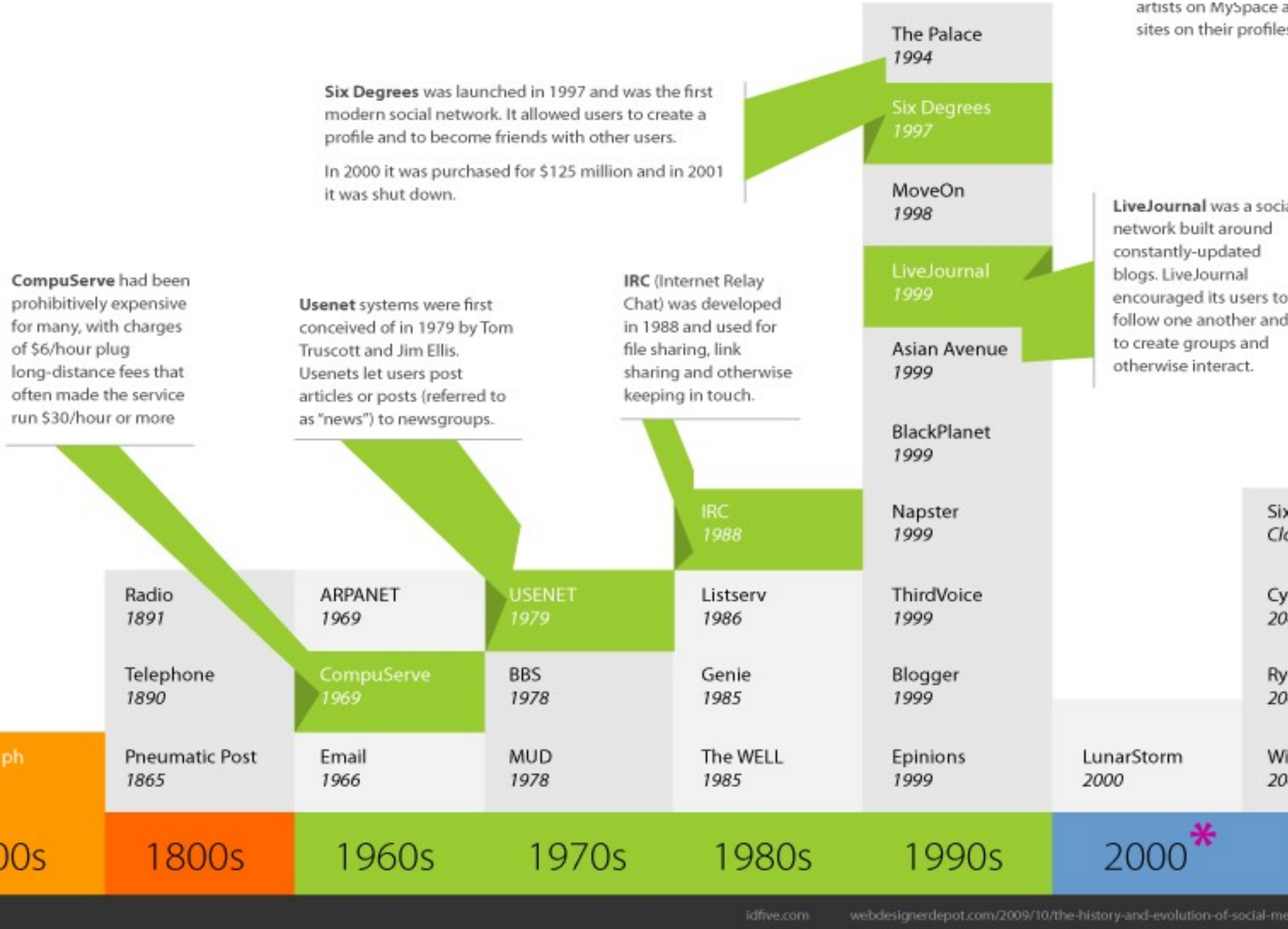
Sample 10.



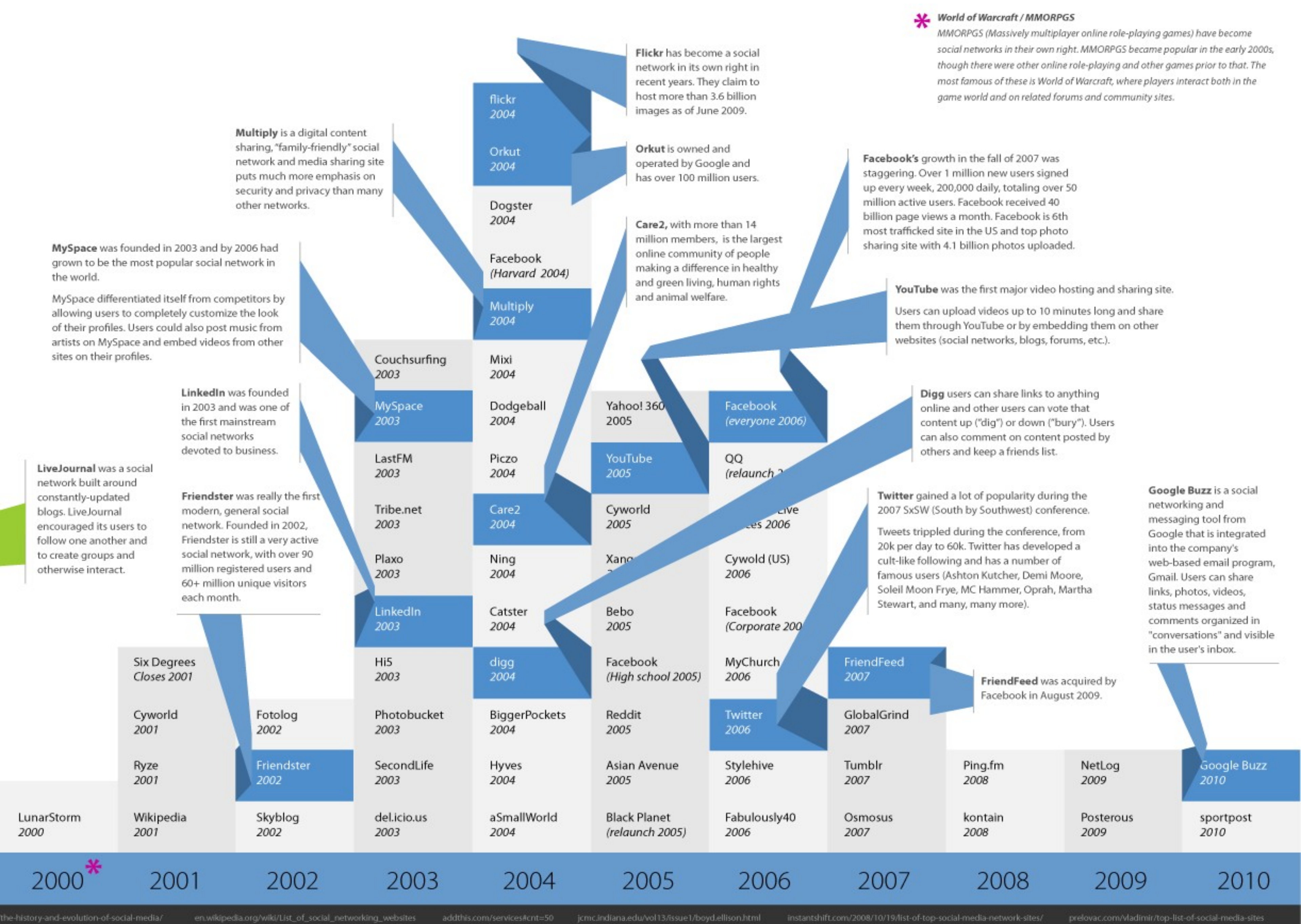
Network representation of brain connectivity: Dorsal and lateral views of the connectivity backbone of human brain. Labels indicating anatomical subregions are placed at their respective centers of mass. Nodes (individual ROIs) are coded according to strength and edges are coded according to connection weight.

How Long They've Been Around?





Source: http://dustn.tv/stay-on-the-cutting-edge/



Source: http://dustn.tv/stay-on-the-cutting-edge/

Why Should We Study Them?

- Networks provide powerful ways of looking at complex data and systems:
  - Spread of news or diseases
  - Evolution of science
  - Structure of the Web
  - Markets & models of trades

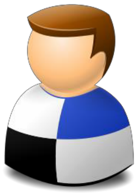
Cheap and high-resolution views into population behavior!

- Networks help to understand if a principle holds across many settings and fields, and

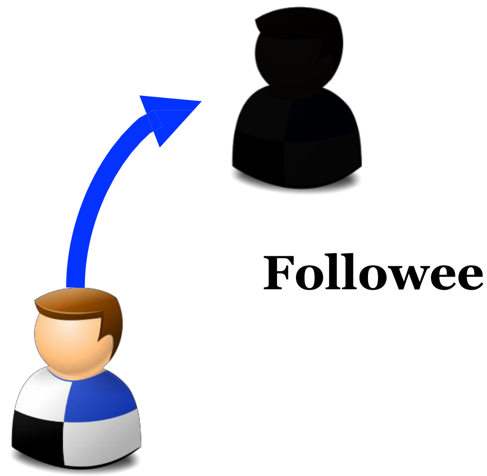
- There are lots of them!

Let's Take a Closer Look at Twitter

- Simple Structure

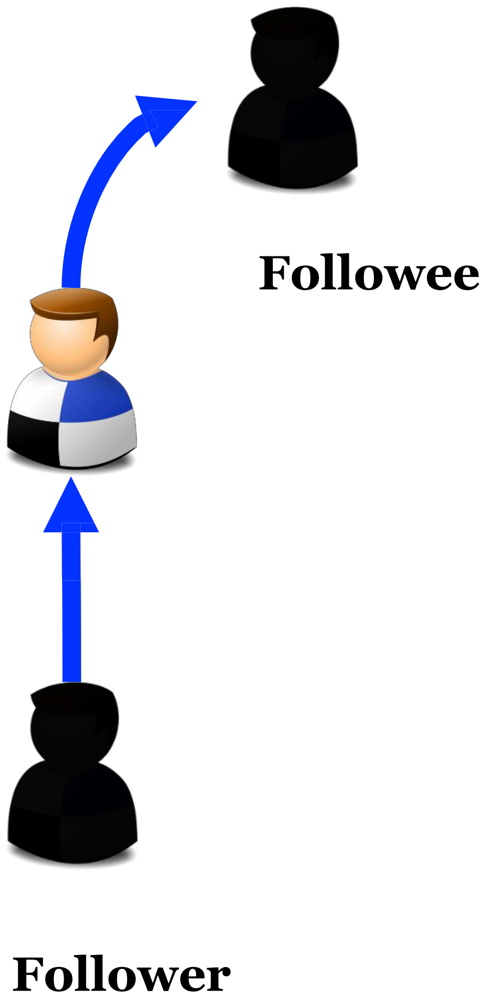


- Simple Structure
- Following
  - To subscribe to other people's posts

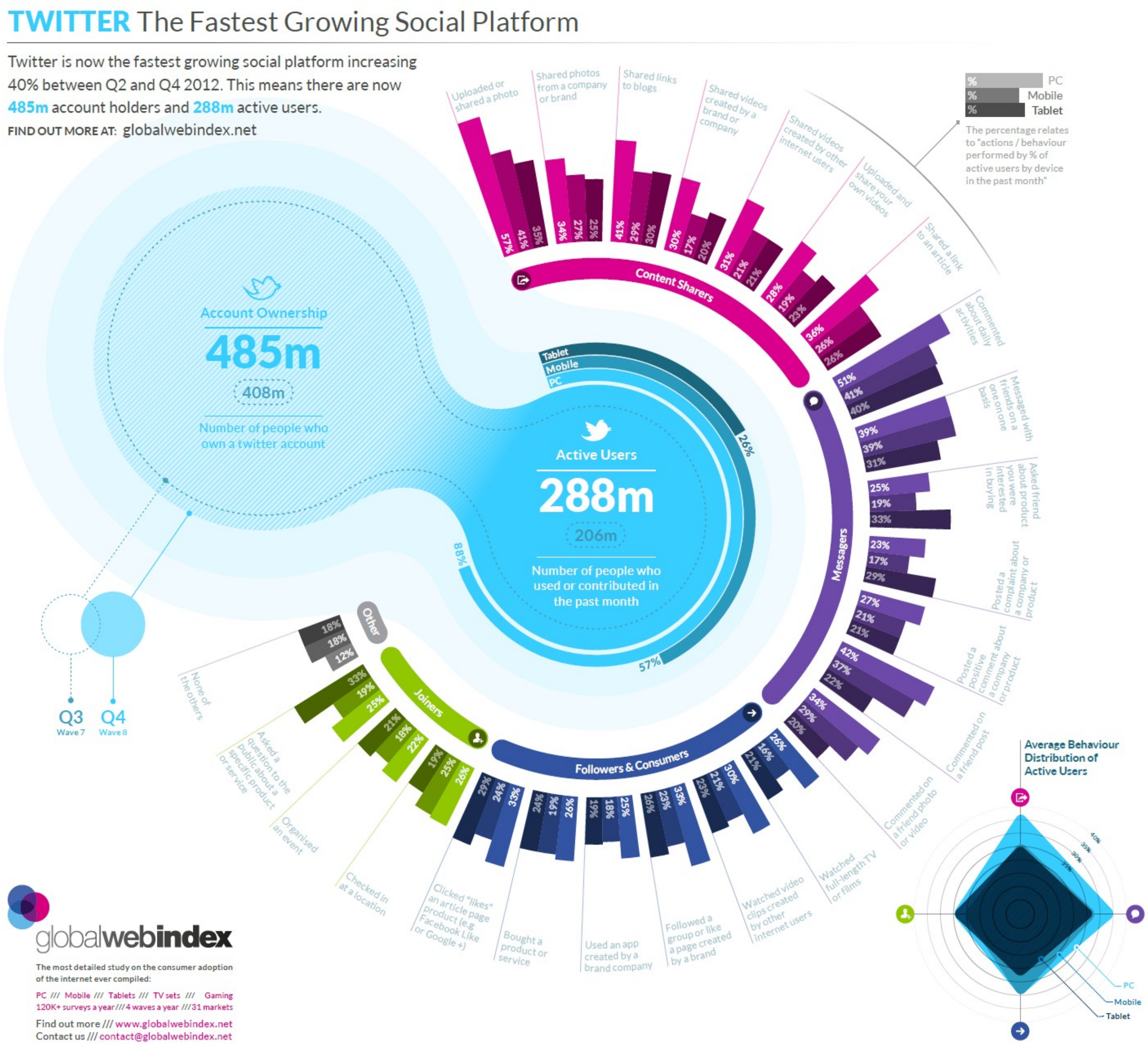


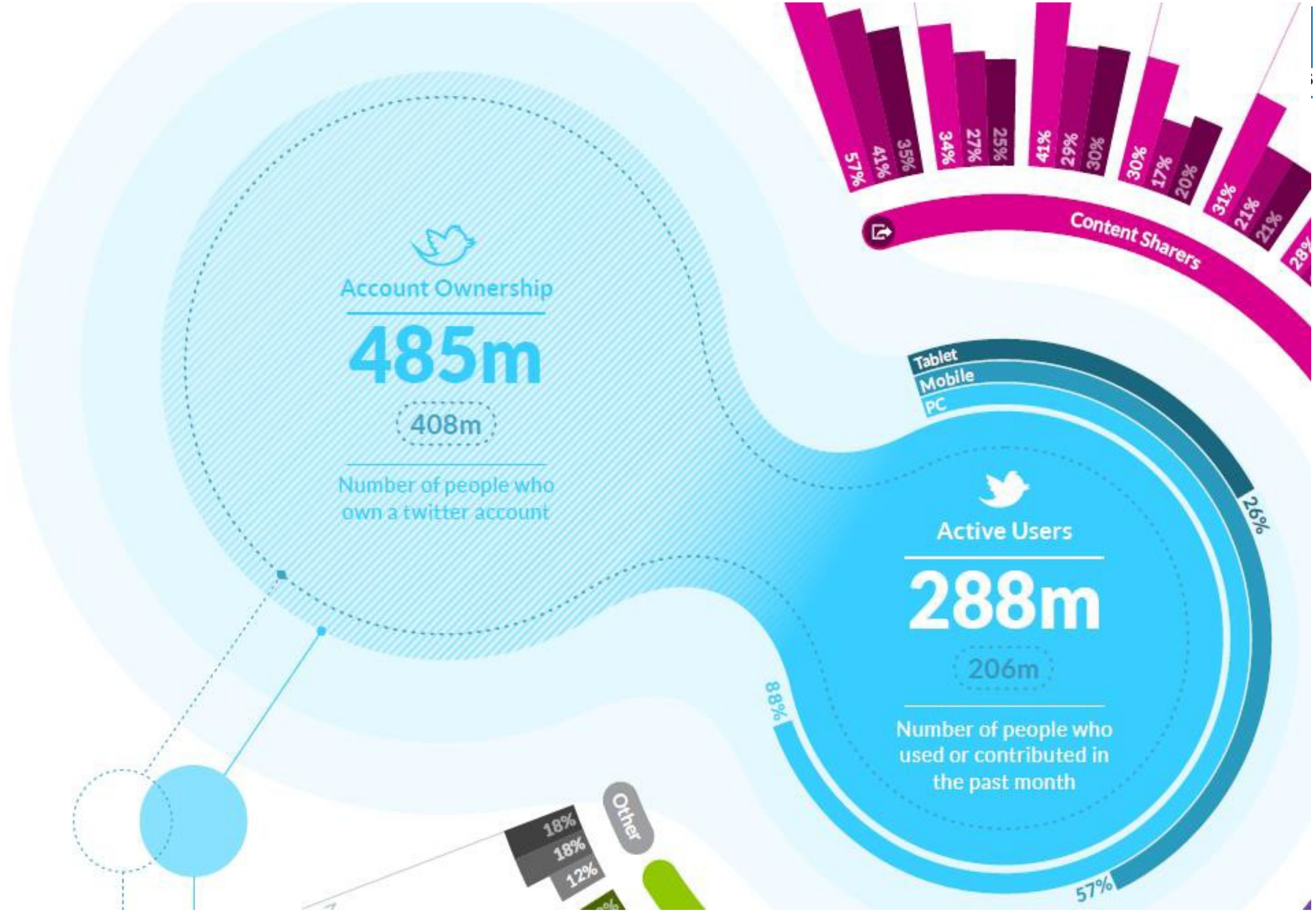
- Simple Structure

- Following
  - To subscribe to other people's posts



UMASS LOWELL

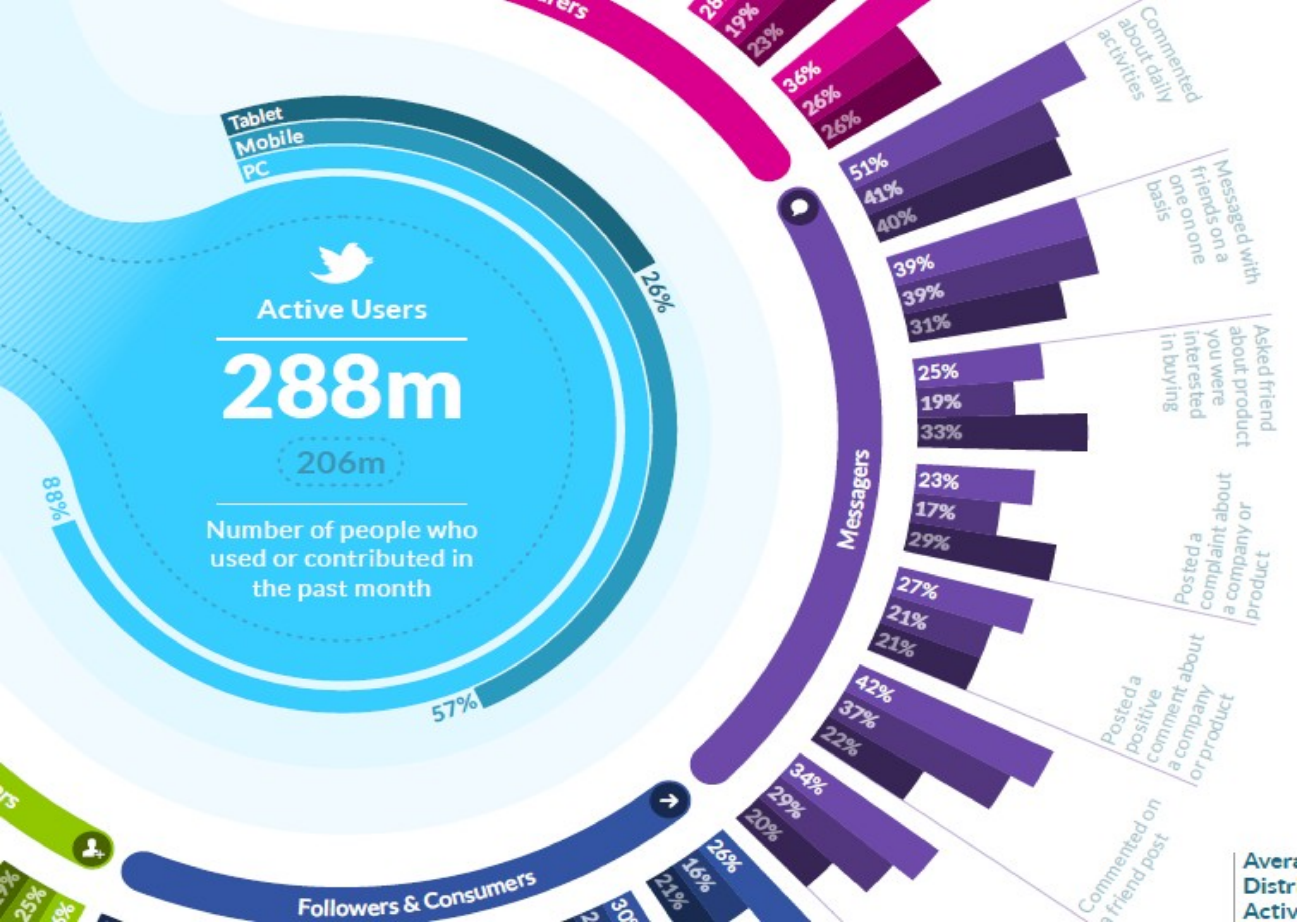




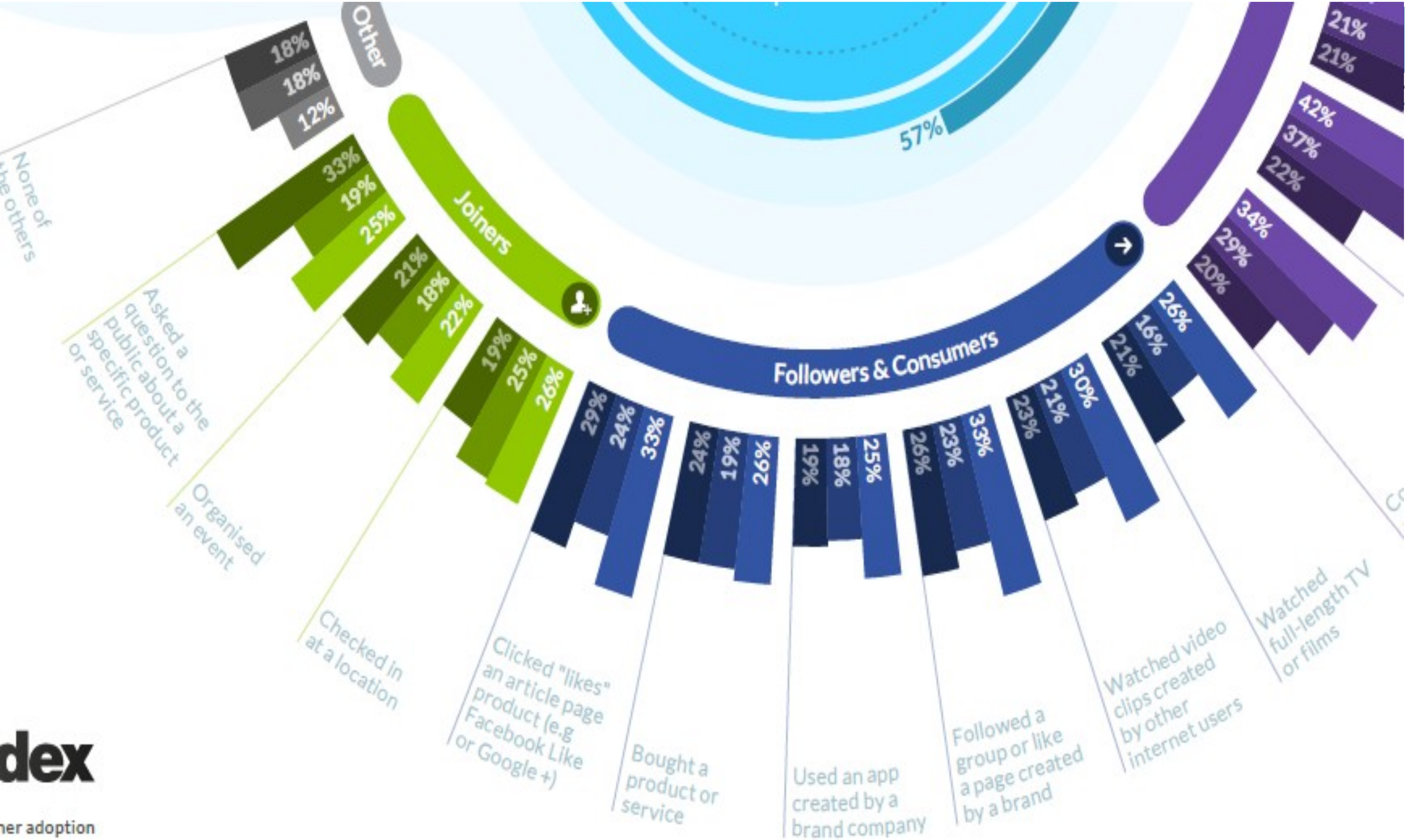
Account Ownership
485m
408m
Number of people who own a twitter account
Active Users
288m
206m
Number of people who used or contributed in the past month
Tablet
Mobile
PC
26%
88%
57%
Content Sharers
57%
41%
35%
34%
27%
25%
41%
29%
30%
30%
17%
20%
31%
21%
21%
Other
18%
18%
12%

al Platform

Uploaded or shared a photo
57%
41%
35%
Shared photos from a company or brand
34%
27%
25%
Shared links to blogs
41%
29%
30%
Shared videos created by a brand or company
30%
17%
20%
Shared videos created by other internet users
31%
21%
21%
Uploaded and share your own videos
28%
19%
23%
Shared a link to an article
36%
26%
26%
Commented about daily activities
51%
Content Sharers
Tablet
Mobile
PC
% PC
% Mobile
% Tablet
The percentage relates to "actions / behaviour performed by % of active users by device in the past month"

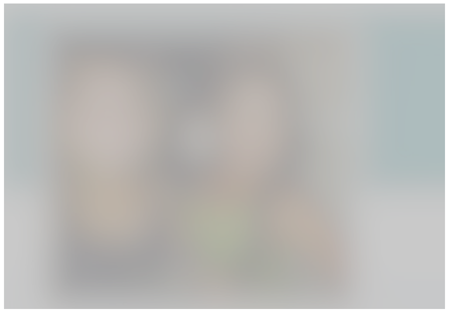


Active Users
288m
206m
Number of people who used or contributed in the past month
Tablet
Mobile
PC
26%
88%
57%
Messagers
Followers & Consumers
Commented about daily activities
51%
41%
40%
Messaged with friends on a one on one basis
39%
39%
31%
Asked friend about product you were interested in buying
25%
19%
33%
Posted a complaint about a company or product
23%
17%
29%
Posted a positive comment about a company or product
27%
21%
21%
Commented on a friend post
42%
37%
22%
34%
29%
20%
36%
26%
26%
19%
23%
26%
16%
21%
25%



dex

ner adoption



…
…
Joined July 2010

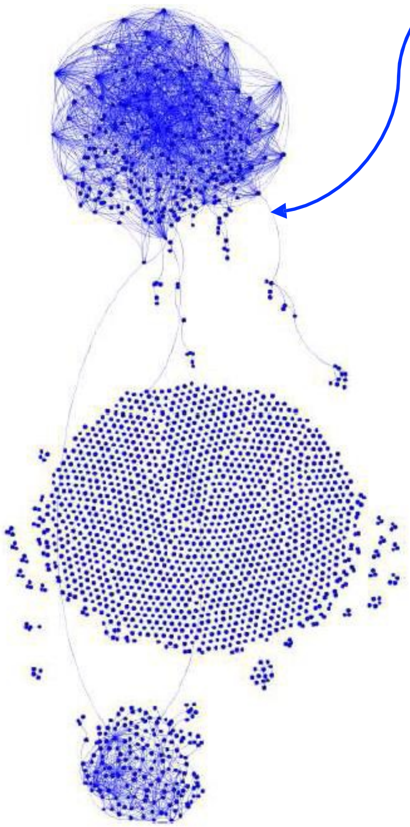
TWEETS	PHOTOS/VIDEOS	FOLLOWING	FOLLOWERS	FAVORITES
477K	215	600	1,219	368

object {21}
created_at : Thu May 01 18:01:19 +0000 2014
id : 461928366862376960
id_str : 461928366862376960
text : Debating if I should switch services with my family or if I should just stay on my own because I reallyyyyy don't want to leave Verizon..
truncated : false
in_reply_to_status_id : null
in_reply_to_status_id_str : null
in_reply_to_user_id : null
in_reply_to_user_id_str : null
in_reply_to_screen_name : null
user {40}
geo : null
coordinates : null
place : null
contributors : null
retweet_count : 0
favorite_count : 0
entities {4}
hashtags [0]
symbols [0]
urls [0]
user_mentions [0]
favorited : false
retweeted : false
lang : en

Net & Content Interactions

User node

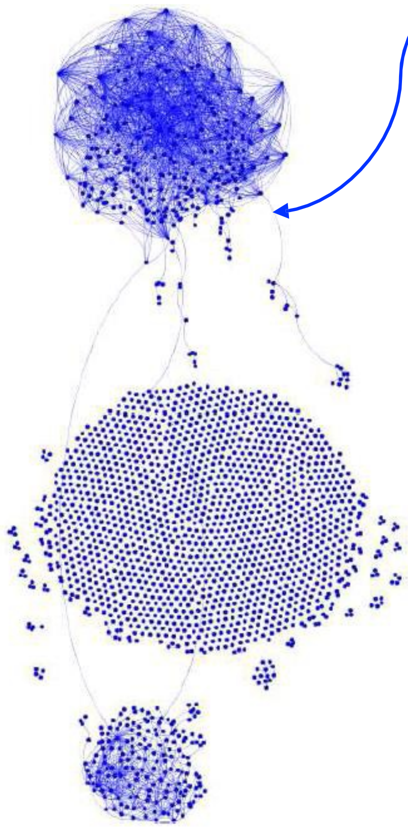
social relations btw users



Net & Content Interactions

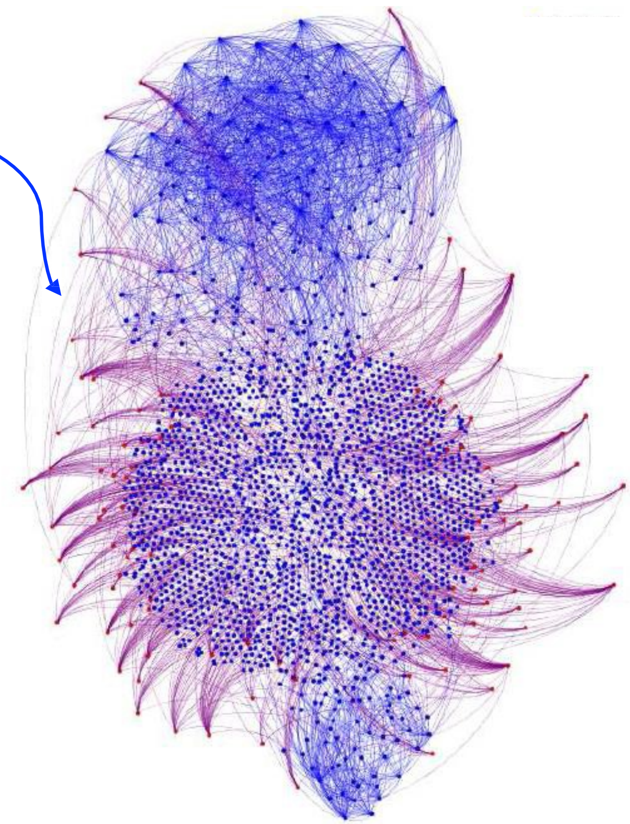
User node

Topic Node



social relations btw users

topical relations btw users and topics



Network Characteristics

- Structure:
  - Network relations are often changing,
  - Weak/strong ties,
  - Often large but still a small world,
  - Popularity dynamics,
  - Cascades, etc.

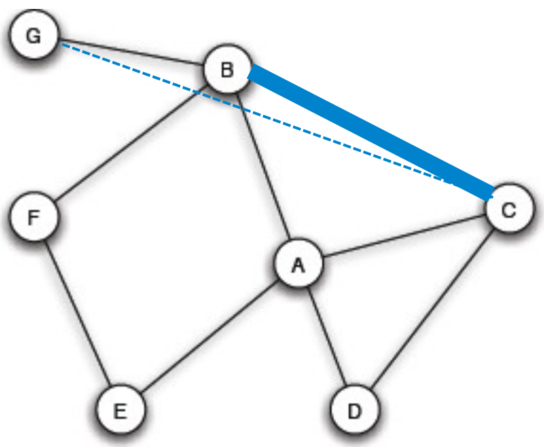
- Content:
  - Streaming type,
  - High prevelance of user-generated/urban words,
  - Often short, context-less, and very noisy, and
  - Various languages.

What Do We Learn?

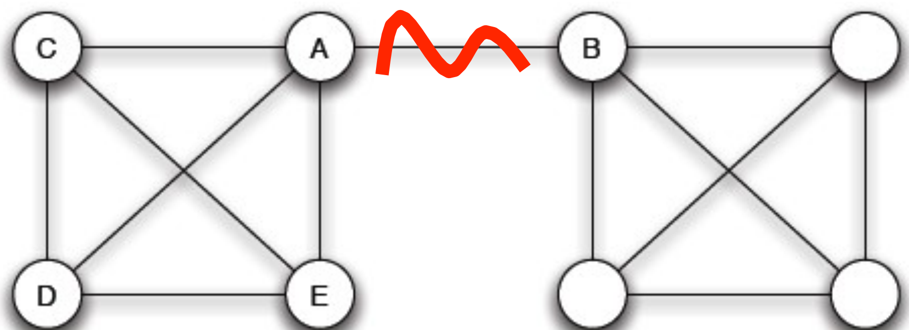
- Graph properties and features
- Node representation
- Graph representation
- Link prediction
- Cascade prediction
- Power laws and Popularity
- Meta Learning with graphs
- Applications
  - Language Analysis
  - Health Informatics
  - Search & Moment Retrieval
  - Trend Detection and Tracking, etc.

What Do We Learn? Cnt.

- Graph Features
  - Strong and Weak Ties



C-B is more likely to form or C-G?



Which link provides access to parts of the net that are unreachable by other means?

Are some nodes more important due to their position in networks?

What Do We Learn? Cnt.

- Graph Features
  - Distance metrics

common neighbors	$\lvert\Gamma(x) \cap \Gamma(y)\rvert$
Jaccard's coefficient	$\frac{\lvert\Gamma(x)\cap\Gamma(y)\rvert}{\lvert\Gamma(x)\cup\Gamma(y)\rvert}$
Adamic/Adar	$\sum_{z\in\Gamma(x)\cap\Gamma(y)} \frac{1}{\log\lvert\Gamma(z)\rvert}$
preferential attachment	$\lvert\Gamma(x)\rvert \cdot \lvert\Gamma(y)\rvert$
Katz$_\beta$	$\sum_{\ell=1}^{\infty} \beta^\ell \cdot \lvert\mathsf{paths}_{x,y}^{\langle\ell\rangle}\rvert$ where $\mathsf{paths}_{x,y}^{\langle\ell\rangle} :=$ {paths of length exactly ℓ from x to y} weighted: $\mathsf{paths}_{x,y}^{\langle 1\rangle} :=$ number of collaborations between x, y. unweighted: $\mathsf{paths}_{x,y}^{\langle 1\rangle} := 1$ iff x and y collaborate.
hitting time stationary-normed commute time stationary-normed	$-H_{x,y}$ $-H_{x,y} \cdot \pi_y$ $-(H_{x,y} + H_{y,x})$ $-(H_{x,y} \cdot \pi_y + H_{y,x} \cdot \pi_x)$ where $H_{x,y}$:= expected time for random walk from x to reach y π_y := stationary-distribution weight of y (proportion of time the random walk is at node y)
rooted PageRank$_\alpha$	stationary distribution weight of y under the following random walk: with probability α, jump to x. with probability $1-\alpha$, go to random neighbor of current node.
SimRank$_\gamma$	$\begin{cases} 1 & \text{if } x = y \\ \gamma \cdot \frac{\sum_{a\in\Gamma(x)}\sum_{b\in\Gamma(y)} \mathsf{score}(a,b)}{\lvert\Gamma(x)\rvert\cdot\lvert\Gamma(y)\rvert} & \text{otherwise} \end{cases}$

What Do We Learn? Cnt.

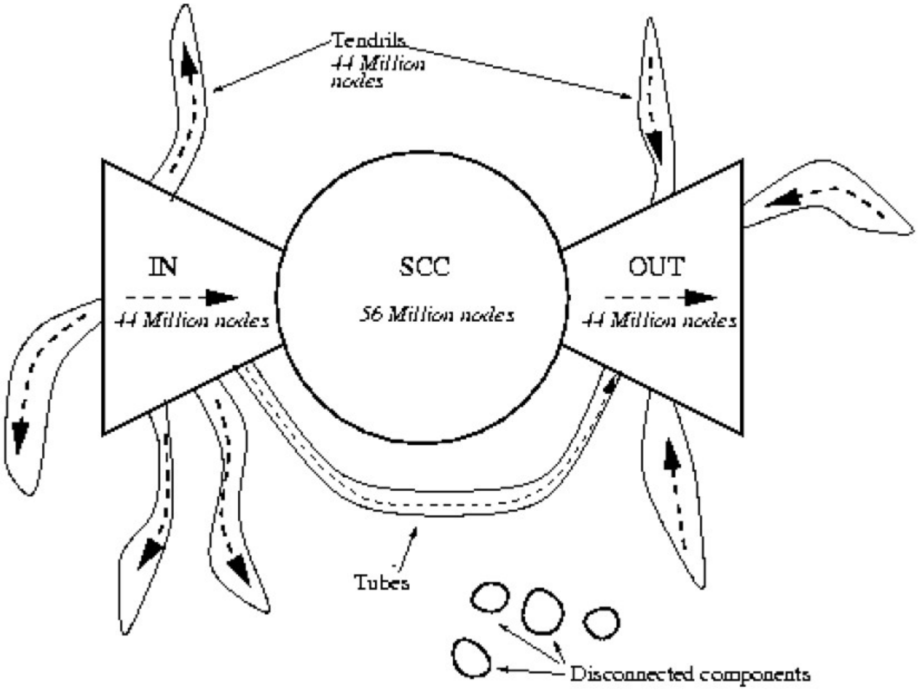
- Graph Features
  - The Structure of the Web
    - ▫ The Web contains a giant SCC

IN nodes:
can reach SCC but cannot be reached from it.

OUT nodes:
can be reached from SCC but cannot reach it.

Tendrils nodes:
(a) reachable from IN but cannot reach SCC,
(b) can reach OUT but cannot be reached from SCC.
Tendrils nodes satisfying both (a) and (b), travel in "tube" from IN to OUT without touching SCC.

Disconnected nodes:
have no path to SCC ignoring directions



99.91% of individuals on FB belong to a single giant connected component

What Do We Learn? Cnt.

- Graph Features
  - Small World Phenomenon

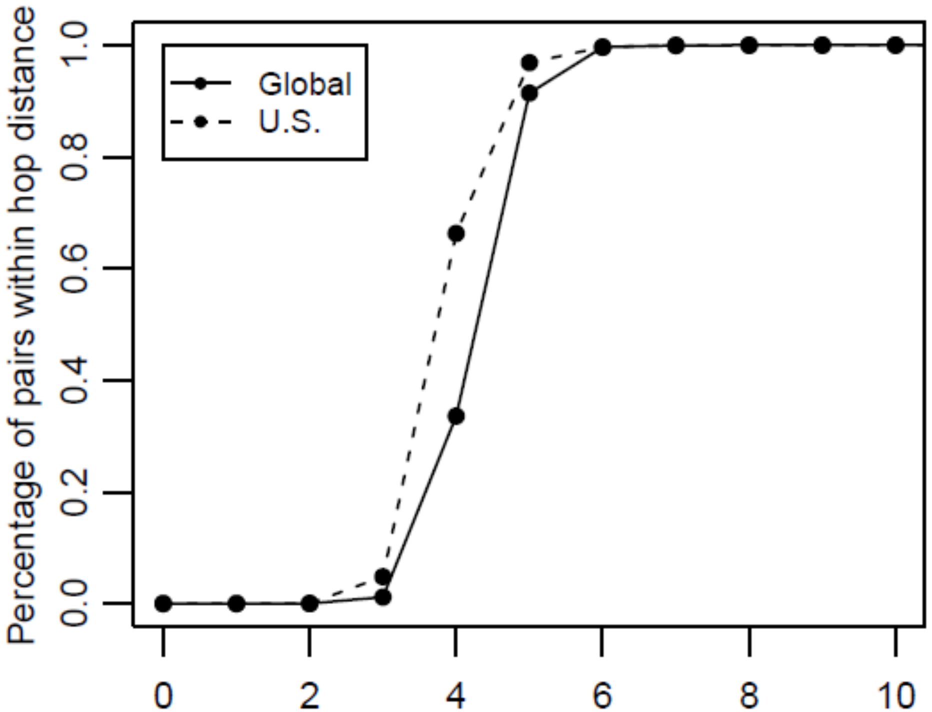


Figure 2. Diameter. The neighborhood function $N(h)$ showing the percentage of user pairs that are within h hops of each other. The average distance between users on Facebook in May 2011 was 4.7, while the average distance within the U.S. at the same time was 4.3.

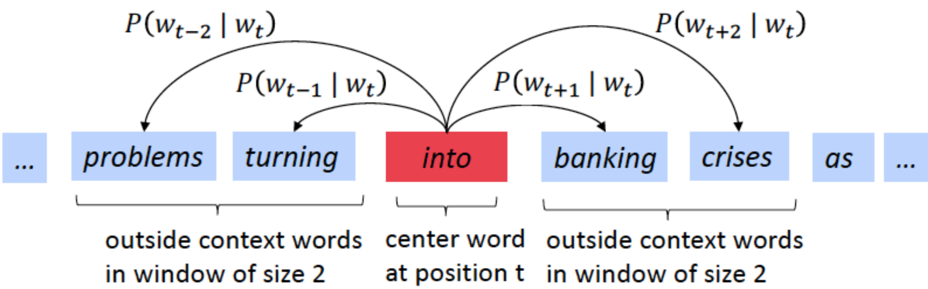
Global
92.0%: within 5 degrees,
99.6%: within six degrees.

U.S. only
96.0%: within 5 degrees,
99.7%: within six degrees.

The Anatomy of the Facebook Social Graph, Ugander, et al., 2011. arxiv.org/abs/1111.4503

What Do We Learn? Cnt.

- Node Representation



$$P(o|c) = \frac{\exp(u_o^T v_c)}{\sum_{w \in V} \exp(u_w^T v_c)}$$

- Update vectors so you can predict well

Nearest words to frog:

1. frogs
2. toad
3. litoria
4. leptodactylidae
5. rana
6. lizard
7. eleutherodactylus

litoria

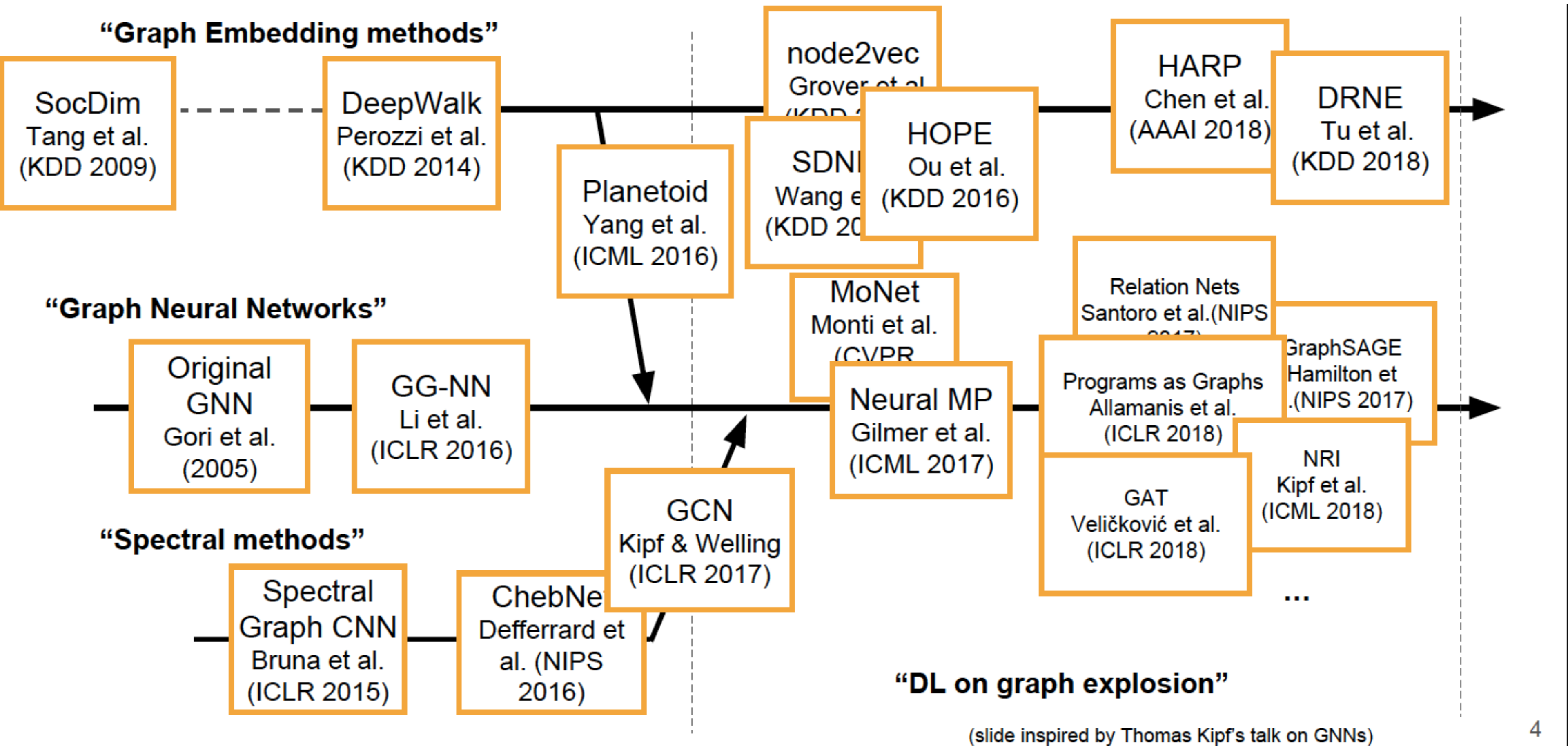
leptodactylidae

rana

eleutherodactylus

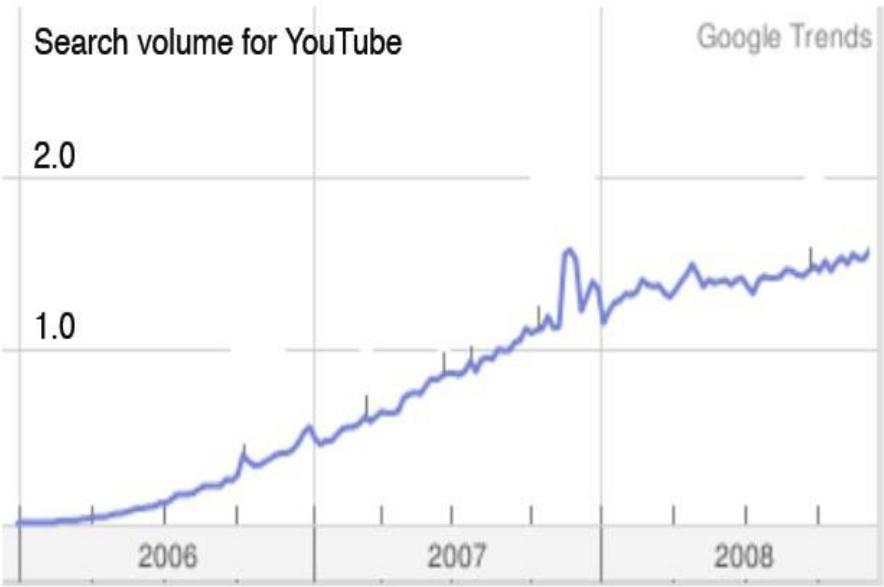
What Do We Learn? Cnt.

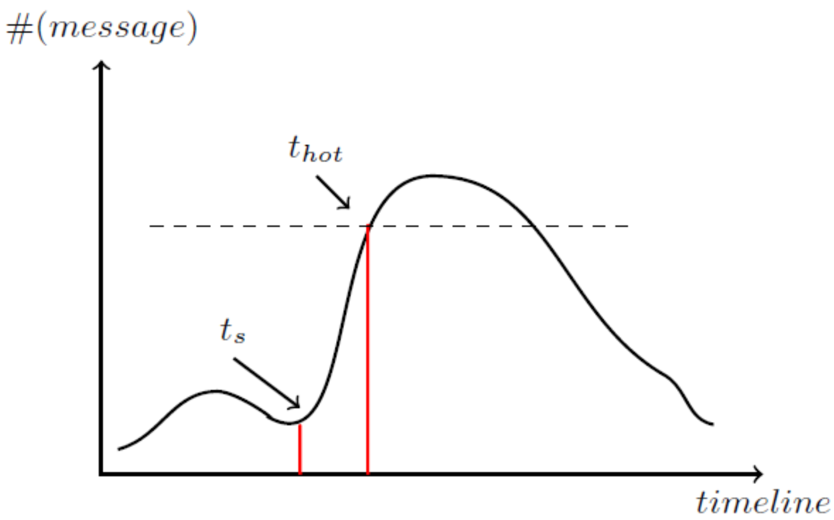
- Graph Representation



What Do We Learn? Cnt.

- Popularity prediction in networks

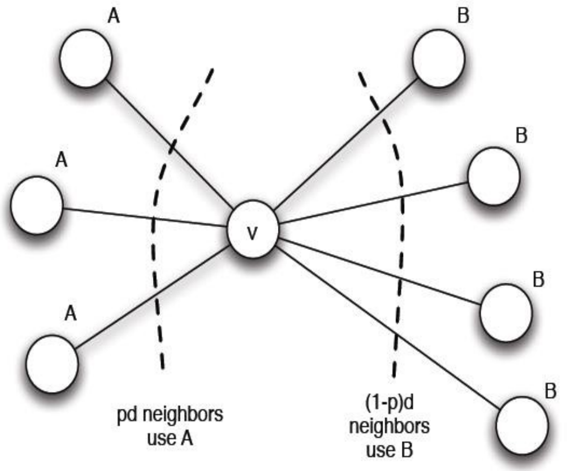


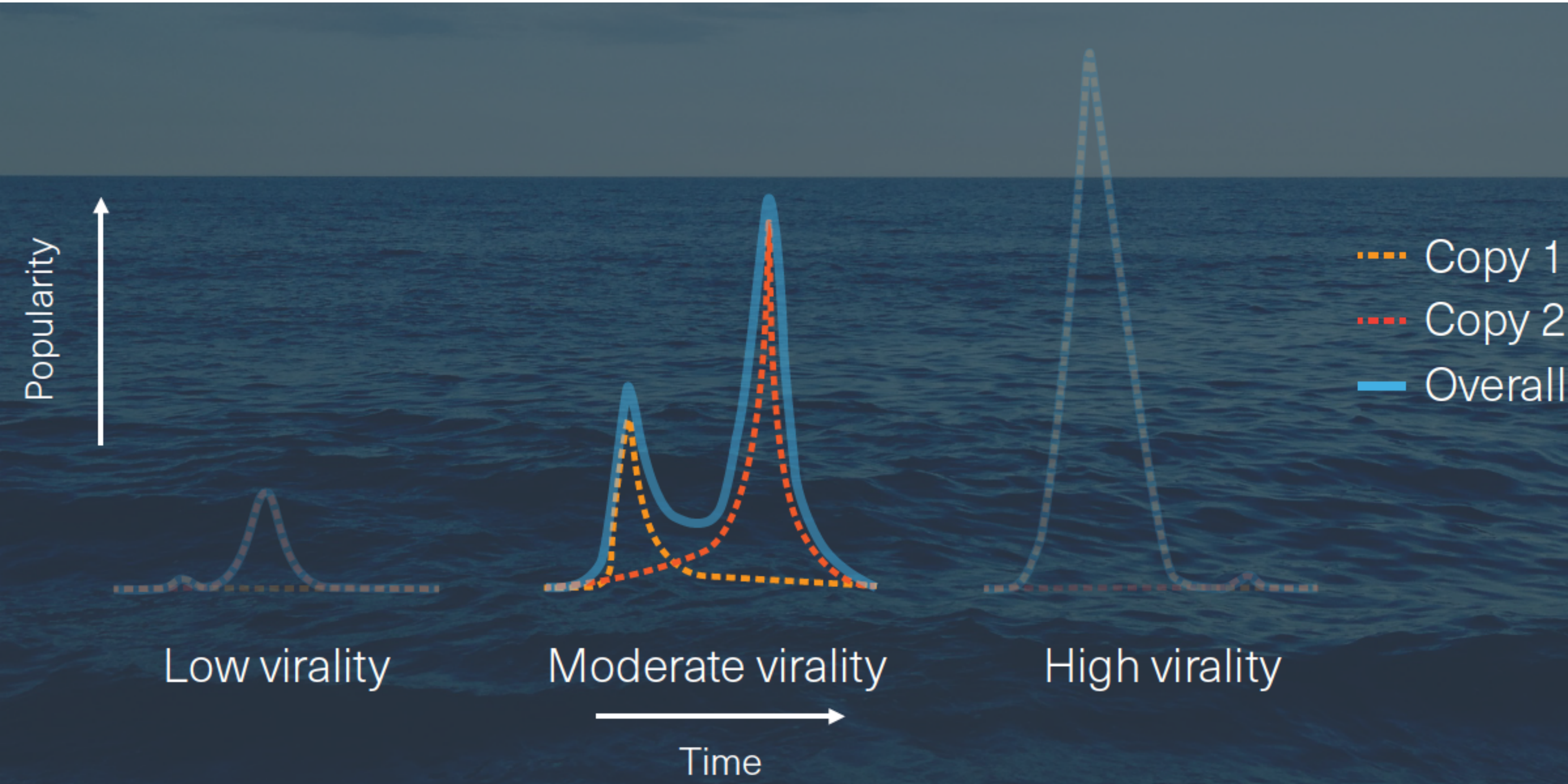


Is it that the rich always get richer? new ideas always get attention and become viral?

What Do We Learn? Cnt.

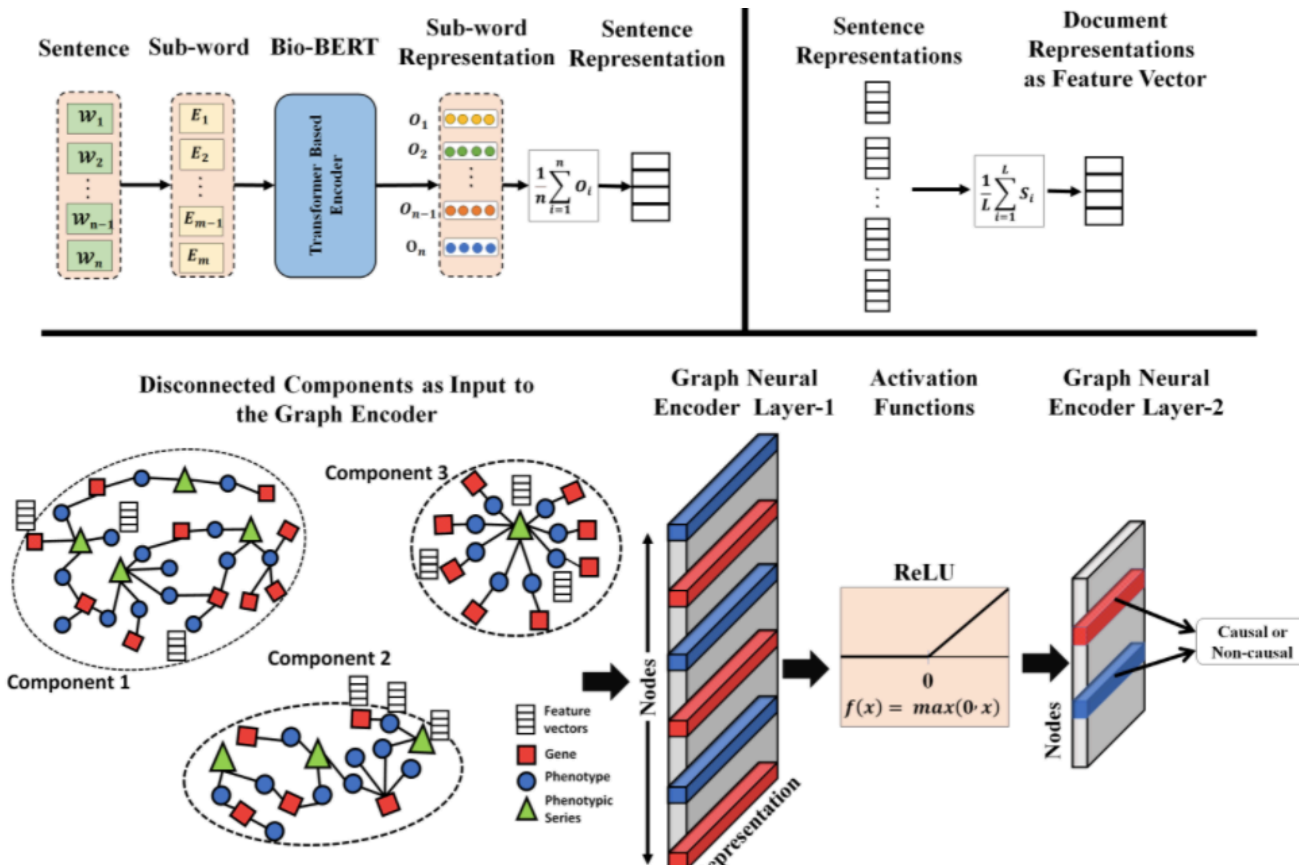
- Cascade Prediction





What Do We Learn? Cnt.

- Link Prediction
  - How can we predict links in networks?

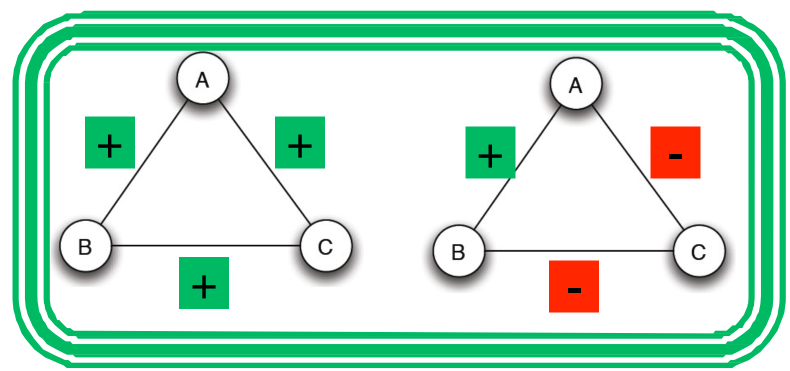


What Do We Learn? Cnt.

- Link Prediction
  - Take a network and annotate its links with
    - + sign representing friendship, and
    - - sign representing antagonism
  - How should we reason about such networks?
    - Say to understand the *tension* between these two forces!

Support / **Oppose** relations

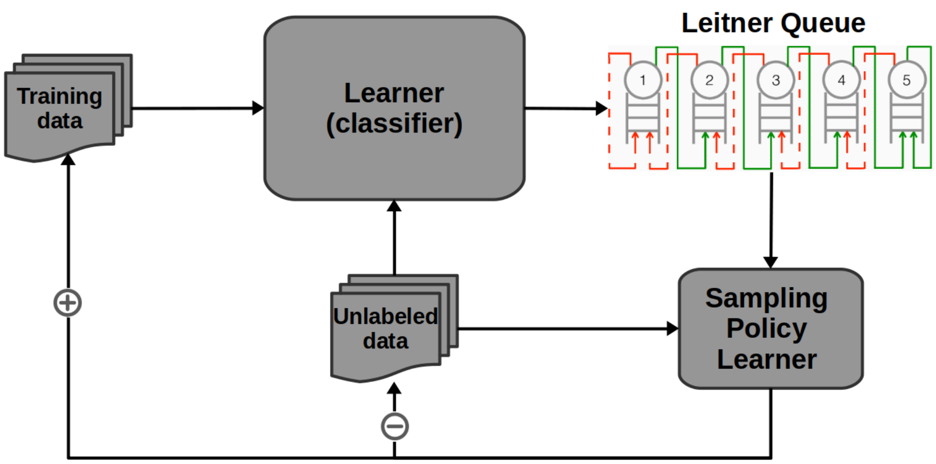
Trust / **Distrust** relations

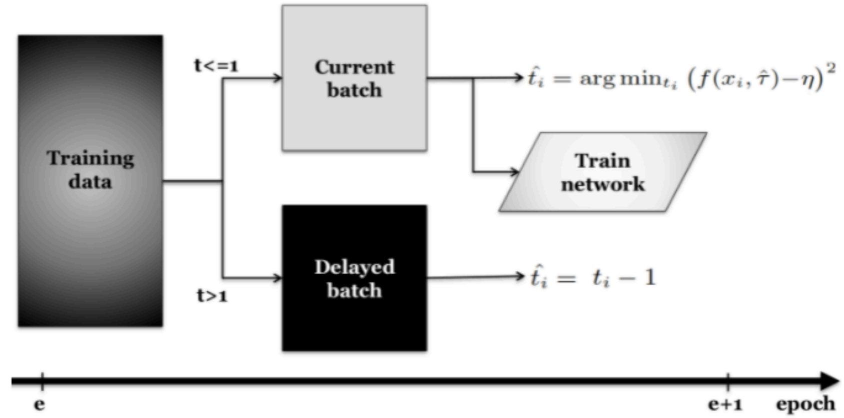


Balanced

What Do We Learn? Cnt.

- Meta Learning
  - Spaced repetition for training
  - Spotting spurious data
  - Neural self-training





Meta Learning

- Meta Learning
  - Curriculum learning with graphs

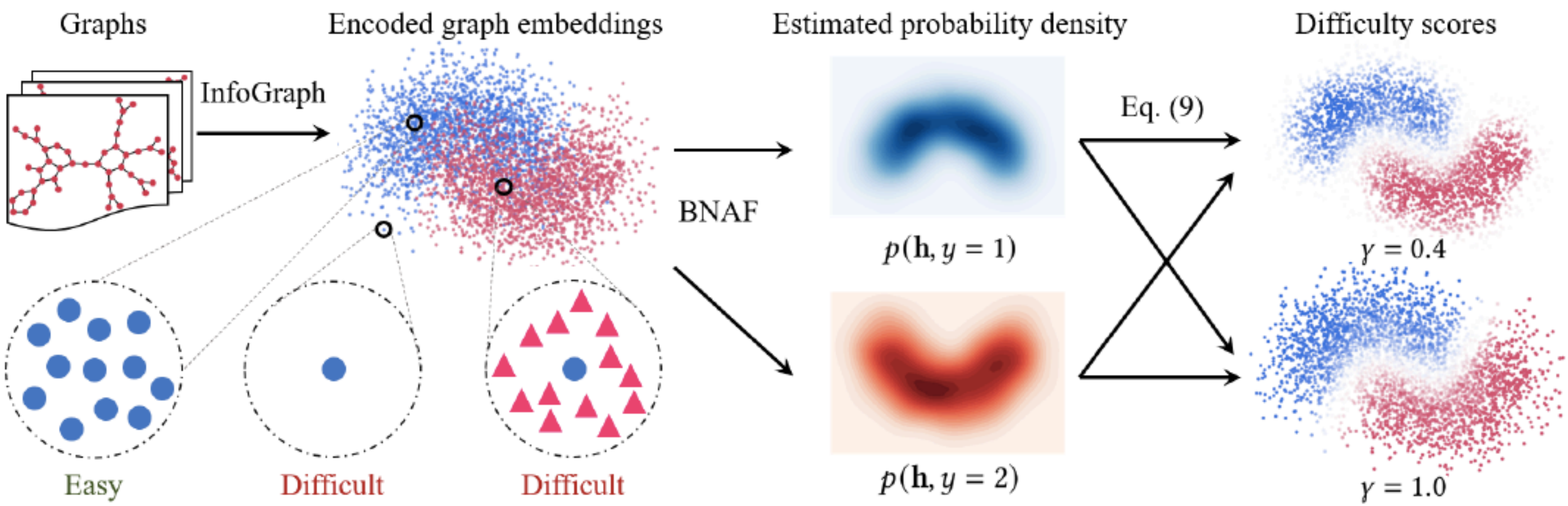


Figure 1: *Infomax Curriculum Design.* We use InfoGraph [49] to obtain graph representations, and BNAF [12] for density estimation. We calculate difficulty scores from intra-class and inter-class densities of Graph Embeddings (by Eq. (9)). Levels of transparency are positively related to difficulty scores. $\gamma = 0.4$ assigns higher difficulty values to outliers than $\gamma = 1.0$.

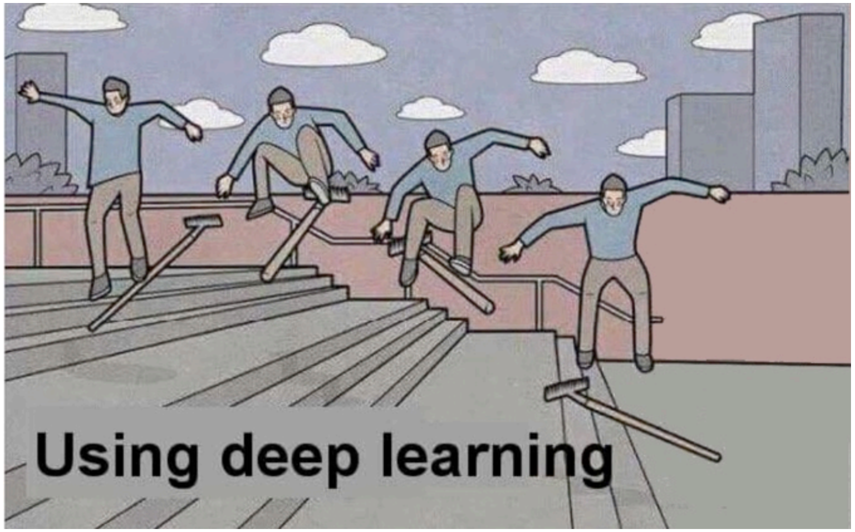
What Do We Learn? Cnt.

Time permitting

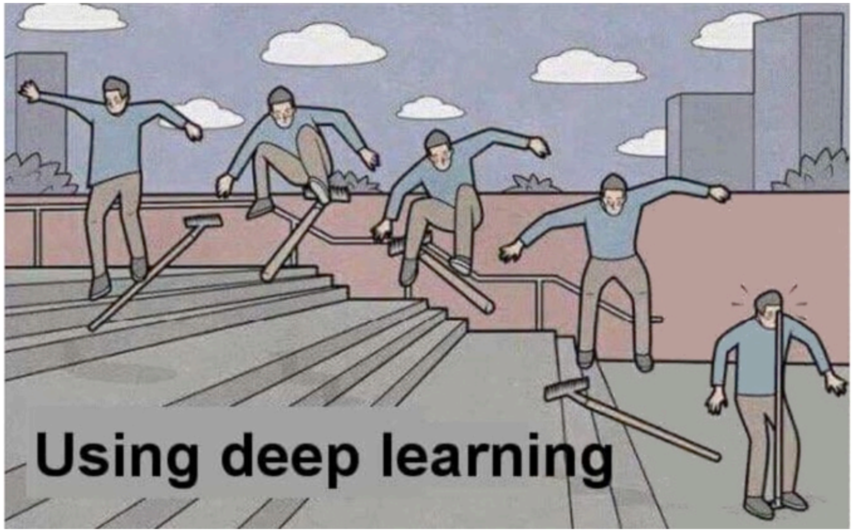
- Applications (*mainly given guest lectures*)
  - Health Informatics
  - Search and Factuality
  - Topic Detection and Tracking

Figure 1: Temporal video moment localization is designed to localize a moment (the red bar) with a start point (24th s) and an end point (30th s) in the video according to the given language query. Here the green bar denotes the ground truth, the orange bar stands for the result of sliding window moment retrieval, and the red bar refers to the localizing result.


Using traditional machine learning methods


Using deep learning


Using traditional machine
learning methods


Using deep learning

Reading

- Ch.01 Introduction [GRL]
- Ch.01 Overview [NCM]