

社交网络与舆情分析

计算机科学与技术学院 任昭春

Outline

- 个人简介
- 课程简介
- 课程设置与安排
 - 课程目标
 - 背景知识
 - 数据与code

Short Break (10min)

- Lecture1: Introduction to Social Media Mining

Outline

- 个人简介
- 课程简介
- 课程设置与安排
 - 课程目标
 - 背景知识
 - 数据与code

Short Break (10min)

- Lecture1: Introduction to Social Media Mining

个人介绍

- 2012 - 2016: PhD in University of Amsterdam, Amsterdam, the Netherlands, Supervisor: Prof. Dr. Maarten de Rijke
- 2012 : Visiting Scholar in Max-Planck-Institut für Informatik, Saarbrücken, Germany.
- 2016 - 2017: Research Associate in University College London, supervised by Prof. Ingemar Cox.
- 2017 - 2019: Research Scientist in the Data Science Lab at JD.com, China.
- 2018 - 2019: Senior Research Manager at JD.com, China.
- 2019– present: Professor at Shandong University

Research interests

- Information Retrieval
- Content analysis in social media
- Semantic search in e-discovery
- Search results diversification
- Recommender systems
- Dialogue systems
- Question-answering
- Summarization

My google scholar page...



Zhaochun Ren 

Professor, [Shandong University](#)
Verified email at sdu.edu.cn - [Homepage](#)

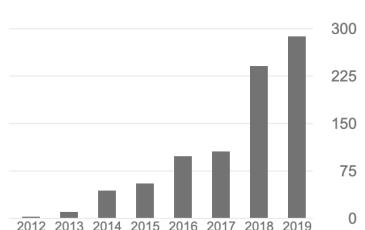
Information retrieval Natural language processing Social media mining

[FOLLOW](#)

	CITED BY	YEAR
Neural attentive session-based recommendation J Li, P Ren, Z Chen, Z Ren, T Lian, J Ma Proceedings of the 2017 ACM on Conference on Information and Knowledge ...	91	2017
Personalized time-aware tweets summarization Z Ren, S Liang, E Meij, M de Rijke Proceedings of the 36th international ACM SIGIR conference on Research and ...	77	2013
Fusion helps diversification S Liang, Z Ren, M De Rijke Proceedings of the 37th international ACM SIGIR conference on Research ...	51	2014
Social Collaborative Viewpoint Regression with Explainable Recommendations Z Ren, S Liang, P Li, S Wang, M de Rijke The 10th ACM International Conference on Web Search and Data Mining, 10	47	2017
Neural rating regression with abstractive tips generation for recommendation P Li, Z Wang, Z Ren, L Bing, W Lam Proceedings of the 40th International ACM SIGIR conference on Research and ...	42	2017
Hierarchical Multi-Label Classification of Social Text Streams Z Ren, MH Peetz, S Liang, W van Dolen, M de Rijke Proceedings of the 37th international ACM SIGIR conference on Research ...	41	2014
Leveraging contextual sentence relations for extractive summarization using a neural attention model P Ren, Z Chen, Z Ren, F Wei, J Ma, M de Rijke Proceedings of the 40th International ACM SIGIR Conference on Research and ...	39	2017
Personalized search result diversification via structured learning S Liang, Z Ren, M De Rijke	36	2014

Cited by

	All	Since 2014
Citations	851	833
h-index	16	16
i10-index	27	24



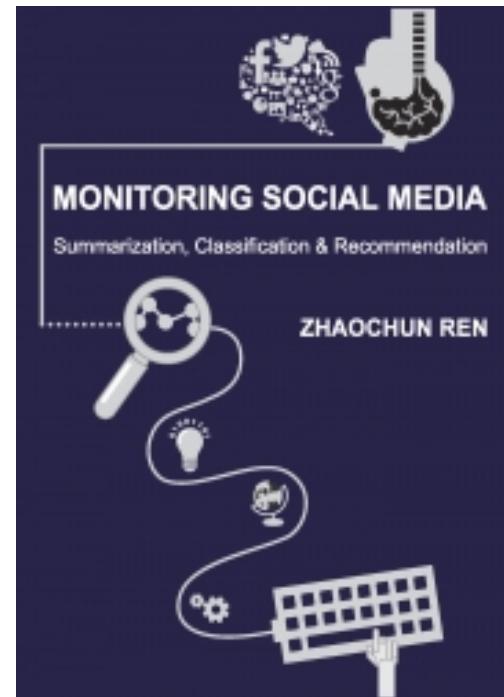
The chart shows a significant increase in citations over time, starting around 2013 and reaching approximately 300 citations in 2019.

Co-authors

	EDIT
 Maarten de Rijke University of Amsterdam	>
 Shangsong Liang Sun Yat-sen University	>
 Dawei Yin Senior Director at JD.com	>
 Pengjie Ren University of Amsterdam	>
 Piji Li Tencent AI Lab	>
 Jiliang Tang Michigan State University	>

Monitoring social media

Summarization, classification and recommendation



Outline

- 个人简介
- 课程简介
- 课程设置与安排
 - 课程目标
 - 背景知识
 - 数据

Short Break (10min)

- Lecture1: Introduction to Social Media Mining

社交网络与社交媒体

- 社交网络
 - 社交网络即社交网络服务，源自英文SNS（Social Network Service）的翻译，中文直译为社交网络服务，意译为社交网络服务。社交网络含义包括硬件、软件、服务及应用，由于四字构成的词组更符合中国人的构词习惯，因此人们习惯上用社交网络来代指SNS（Social Network Service）。
- 社交媒体
 - 社交媒体（social media）是人们用来创作、分享、交流意见、观点及经验的虚拟社区和网络平台。社交媒体和一般的社会大众媒体最显著的不同是，让用户享有更多的选择权利和编辑能力，自行集结成某种阅听社群。社交媒体并能够以多种不同的形式来呈现，包括文本、图像、音乐和视频。

社交网络与社交媒体

- 互联网所提供的连通性和数字化，前所未有地促进了以数字为介质的社会交往。
- 数字化交往的普及蕴含着一种风险：使人们脱离真切的群体生活，沦为缺乏现实感和责任心的、沉浸于自我的“网络流民”，造成整个社会结构的散众化。
- “散众社会”作为一种动员结构，蕴含着充沛的集体行动潜能，在特定条件下可能演化为网络群体性事件，对社会秩序产生深刻的影响。

社交网络中的社会。。。。

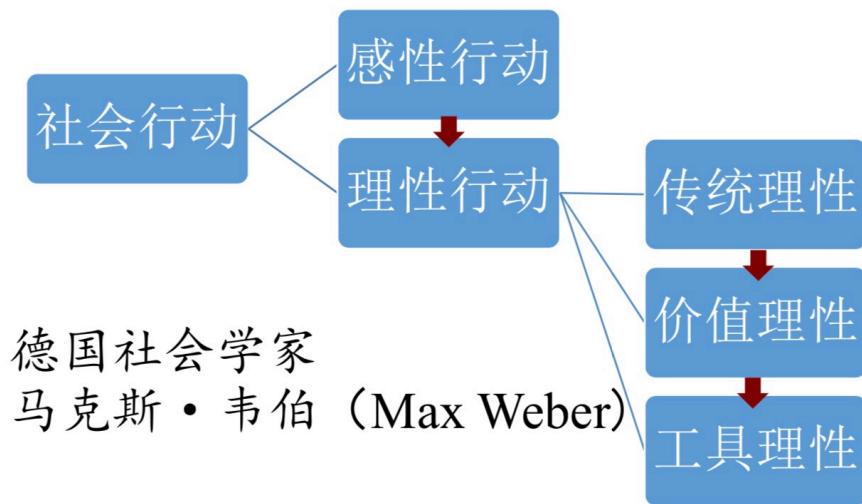
连通性

人与人、人与物之间通路更多、更密、更廉价，万物之间的关联更加畅通。

数字化

事物本身以及事物之间关系的信息越来越多地被以数字的形式记录和构建。

社会行动的逻辑



■使任性和感性的外化更加容易、更加方便

- 情感的激发、撩拨、流布与感染 {段子手、朋友圈、网络主播,}
- 情感的商业化 {抖音、西瓜视频,}
- 感性消费

Facebook

facebook

Search

Home Profile Account

Mark Zuckerberg

Has worked at Facebook Studied Computer Science at Harvard University Lives in Palo Alto, California From Dobbs Ferry, New York Born on May 14, 1984





Education and Work

Employers

 Facebook
Feb 2004 to present - Palo Alto, California
FBX Profile

College

 Harvard University
Computer Science · Psychology
CS182. Intelligent Machines with Andrew Bosworth
CS121. Introduction to Computational Theory with James Wang and Kang-Xing Jin

High School

 Ardsley High School

 Phillips Exeter Academy
Class of 2002

Family

 Karen Zuckerberg
Mother

 Edward Zuckerberg
Father

 Randi Zuckerberg
Sister

 Donna Zuckerberg
Sister

 Arielle Zuckerberg
Sister

You and Mark

Send Message Poke

3 Mutual Friends

Sponsored

 Police Auctions
gsaauctions.gov
Like a slick deal? Now you can get up to 90% retail with police seized auctions. Get in on the action.

 SF Bucket List
partners.livingsocial.com
Things to do in San Francisco before you die. One huge coupon emailed daily.

 Stay close to your team
Check the score and see highlights from the big game with AT&T High Speed Internet for only \$14.95/mo.

 Craft Beer Attorney
Need legal assistance with your California craft beer?

- How does Facebook use your data?
- Where do you think Facebook can use your data?

What about Amazon?

Amazon.com: Online Shopping

www.amazon.com

amazon Your Amazon.com | Today's Deals | Gift Cards | Help

Shop by Department ▾ Search All ▾ Go

Hello, Sign in Your Account ▾ Cart ▾ Wish List ▾

FREE Two-Day Shipping [Join Amazon Prime Today](#)

Amazon Instant Video: The latest TV shows, available instantly. EXPAND TO LEARN MORE

Instant Video MP3 Store Cloud Player Kindle Cloud Drive Appstore for Android Digital Games & Software Audible Audiobooks

The All-New kindle fire HD The ultimate HD experience From \$199 [Shop now](#)

Introducing kindle paperwhite The world's most advanced e-reader From \$119 [Shop now](#)

Bikes with Street Cred Clothing Trends Amazon Prime

THE AMAZON CLOTHING STORE DRESS SHOP

The season's biggest trends from Calvin Klein, BCBGMAXAZRIA, Tracy Reese, and more.

[Shop Dresses](#) [Shop All Clothing](#)

Friends & Family Gifting [Gift Connections Made Easy](#) [Learn more](#)

\$0 Hundreds of Free Songs Play on the Web, Android, and Kindle Fire [Shop now](#) amazon MP3

Advertisement

Gold Box Deal of the Day

\$1.99 Kindle Inspirational Memoirs and More

Yelp?

English ▾

Friends' Activity Sign Up for Yelp Log In

yelp
Real people. Real reviews. ®

Welcome About Me Write a Review

Try Yelp in... Are You Looking For Yelp Berkeley?

- Amsterdam
- Atlanta
- Austin
- Berlin
- Boston
- Chicago
- Dallas
- Denver
- Detroit
- Dublin
- Honolulu
- Houston

San Francisco Community Acupuncture

San Francisco Food Bank

San Francisco Symphony

San Francisco restaurants

San Francisco Movie Tours

San Francisco Marriott Fishermans Wharf

San Francisco Gay Wedding Video

San Francisco Soup Company

San Francisco Test Only Smog

San Francisco CityPASS

Yelp Berkeley

Search for (e.g. taco, cheap dinner, Max's)
san francisco restaurants

Near (Address, Neighborhood, City, State or Zip)
San Francisco, CA

Search

Talk Events Member Search Hide

Philadelphia

Portland

Saint Louis

San Diego

San Francisco

San Jose

Seattle

Toronto

Vancouver

Washington, DC

More Cities »

Yelp is the fun and easy way to find and talk about great (and not so great) local businesses [Sign Up Now](#)

Best of Yelp

[More "Best Of" »](#)

Restaurants
4505 reviewed

1. La Bedaine
2. Kingston 11 Cuisine
3. Vital Vittles
4. Cheese Board Pizza
5. Emilia's Pizzeria
[...see more »](#)

Nightlife
881 reviewed

Shopping
4852 reviewed

1. Chestnut & Vine Floral...
2. Waterside Workshops
3. UniFormal & UniEleganza Tuxedo...
4. Lee's Florist & Nursery
5. Supple Integrative Skin Care
[...see more »](#)

Beauty and Spas
2566 reviewed

Browse by Category

- Restaurants
- Food
- Nightlife
- Shopping
- Beauty and Spas
- Arts & Entertainment
- Event Planning & Services
- Health and Medical
- Active Life
- Hotels & Travel
- Automotive
- Home Services
- Local Services
- Local Flavor
- Pets
- Education

Get the Yelp app on your mobile phone

It's free and helps you find great, local businesses on the go!

[Get it for Free Now](#)

Review of the Day [Archive »](#)

Voted by our members!

Dopo
*Cozy, romantic little neighborhood spot. Nice location right across from good ol' Fenton's. No reservations at this little place, so either come a little

Or Twitter?

Screenshot of a web browser showing the Twitter "Who to follow" page for interests in technology. The page lists various users and organizations to follow, including **guardiantech**, **kaifulee**, **sacca**, **kevinrose**, **arsTechnica**, **lessig**, and **leolaporte**. It also shows recently followed users like **GlobalFundWomen** and **nockee**, and recently viewed profiles like **StartupReport** and **guardiantech**.

Who to follow

View Suggestions **Browse Interests** **Find Friends**

View Technology (60)

Art & Design · Books · Business · Charity · Deals · Entertainment · Family · Fashion · Food & Drink · Funny · Health · Music · News · Politics · Science · Sports · Staff Picks · Staff Picks: NFL Playoffs · Technology · Travel · Twitter

guardiantech Guardian Tech
News and comment from the Guardian's technology team

kaifulee Kai-Fu Lee I am the CEO of Innovation Works (www.innovation-works.com). I used to work at Microsoft, SGI, Apple, and Google. I have written several Chinese books.

sacca Chris Sacca I'm an investor in Twitter & other startups. Learn more at www.lowercaselle.com. Just finished bicycling across the USA and am thinking about my next adventure.

kevinrose Kevin Rose Tech angel investor. Founder of digg.com, wefollow.com. Random ideas, entrepreneur, climber of rocks, video blogger, & tea drinker.

arsTechnica Ars Technica The Art of Technology

lessig Lessig law prof, reformer.

leolaporte Leo Laporte Podcaster, broadcaster and tech pundit. The Tech

You recently followed · [View All →](#)

GlobalFundWomen Global Fund forWomen · Follow
Largest global grantmaker exclusively funding women's rights. Committed to a world of equality & social justice.

nockee Chris · Follow
Ultracyclist and endurance junkie. Did Race Across Oregon and Hoodoo 500 Voyager in 2009. Doing RAW solo in 2011

You recently viewed · [Clear](#)

StartupReport · Follow Kevin Steger

guardiantech · Follow **guardiantech** Guardian Tech

pierre · Follow Pierre Omidyar

spoonen · Follow Sanjay Poonen

cote · Follow **cote** Stratelux

Invite Friends

Not finding who you're looking for? Invite friends to Twitter via email. See what you'll send them.

your friend's email address **Invite**

Lots of people to invite? Separate multiple email addresses with commas.

Related Services

Microsoft · Bing Social's Recommendations for Influential Twitter Users.

LinkedIn · Find your LinkedIn colleagues who use Twitter.

About · Help · Blog · Status · Jobs · Terms · Privacy · Shortcuts
Advertisers · Businesses · Media · Developers · Resources · © 2011 Twitter

微博？微信？知乎？

有什么新鲜事想告诉大家?

点赞! 香港青年自发上街清理涂鸦标语>>视频 热门微博

表情 图片 视频 话题 头条文章 ... 公开 发布

全部 原创 图片 视频 音乐 文章 搜索我关注人的微博

历史的进城 V

3分钟前 来自 历史的进城Android

从整体数据来看，21只富国旗下基金在本年二季度买入兴齐眼药789万股，占总股本的14.73%，占总股本的9.57%。正好卡在“同一基金管理人管理的全部基金持有一家公司发行的证券，不得超过该证券的10%”的红线边。

@历史的进城 V

网友提供的线索，“兴齐眼药”的半年报显示，有21只富国的基金扎堆在这支最牛眼药水里。互联网主题，军工主题，高端装备主题都有，合力把股价推到93。就问你们服不服？

8月29日 07:35 来自 历史的进城Android

兴齐眼药 sz300573

收藏 转发 1 赞

投行新闻

知乎 首页 发现 等你来答 开学第一课 搜索 提问

推荐 关注 热榜

窗户通风与防尘兼得的方法有哪些?

鹏哥门窗科普：可以通过3种方式实现窗户通风与防尘兼具，门窗安装前考虑到这个问题，可以通过改变窗型：（1）内开内倒窗或「微通风内开内倒窗」改良通风；（2）采用「窗新风」，改变通风方式。门窗安装后，可... 阅读全文

▲ 赞同 135 ▾ 收起评论 分享 收藏 感谢 ...

15条评论 切换为时间排序

星期五Friday 2天前

还是会灰尘进来吧。和新风系统没法比

鹏哥门窗科普 (作者) 回复 星期五Friday 2天前

一分钱一分货不是吗

星期五Friday 回复 鹏哥门窗科普 (作者) 2天前

是。新风系统有个最大的问题，梁上要开孔。很郁闷。

远道而来 1天前

哎哟，不错哟

传统的媒体也开始具有社交属性

The New York Times

OPINION | Texas Is a Leader in Mass Shootings. Why Is the Governor Silent?

Related

Death Toll in West Texas Shooting Climbs to 7

Sept. 1, 2019



Richard Parker is the author of "Lone Star Nation: How Texas Will Transform America."

The Times is committed to publishing a diversity of letters to the editor. We'd like to hear what you think about this or any of our articles. Here are some tips. And here's our email:

最热评论



xialei_026 江苏南京

我就看这乱港闹剧到底能持续到什么时候才能大结局。

9月1日18:26

▲ 赞46 ▲ 回复

[opinion](#)) and



nb伤心人 浙江宁波

警告哈，直接用麻醉枪击倒，然后逮捕入狱。

9月1日18:25

▲ 赞36 ▲ 回复



会计通 山东济南

别几把天天报了，又解决不了

9月1日21:23

▲ 赞13 ▲ 回复

最新评论



安徽有山啊 上海

怎么到9月份，还没有控制住。

51分钟前

▲ 赞4 ▲ 回复



猛虎和蔷薇啊 广东广州

快了，这几天冲突升级，这是最后的疯狂，我猜最多一星期。

9月1日23:47

▲ 赞7 ▲ 回复



会计通 山东济南

别几把天天报了，又解决不了

9月1日21:23

▲ 赞13 ▲ 回复

查看全部11条评论 >

Outline

- 个人简介
- 课程简介
- 课程设置与安排
 - 课程目标
 - 背景知识
 - 数据与code

Short Break (10min)

- Lecture1: Introduction to Social Media Mining

课程信息

- **Instructor:**
 - Zhaochun Ren (zhaochun.ren@sdu.edu.cn)
- **Teaching Assistant**
 - Dongdong Li
 - Zhongkun Liu
 - Daoyuan Fu
- **授課时间 (Starting Sep 2nd 2019):**
 - 1-17周 Monday 1-2节 地点：振声苑N115
 - 实验课：2-17周 Tuesday 7-8节 地点：机房101,105,109

Outline

- 个人简介
- 课程简介
- 课程设置与安排
 - 课程目标
 - 背景知识
 - 数据

Short Break (10min)

- Lecture1: Introduction to Social Media Mining

课程目标

- Understand social aspects of the Web
 - Social Theories + Social media + Mining
 - Learn to collect, clean, and represent social media data
 - How to measure important properties of social media and simulate social media models
 - Find and analyze communities in social media
 - Understand how information propagates in social media
 - Understanding friendships in social media, perform recommendations, and analyze behavior
- Study or ask interesting research issues
 - e.g., start-up ideas / research challenges
- Learn representative algorithms and tools

Outline

- 个人简介
- 课程简介
- 课程设置与安排
 - 课程目标
 - 背景知识
 - 数据与code

Short Break (10min)

- Lecture1: Introduction to Social Media Mining

背景知识

- Prerequisites:
 - Data Structures and Algorithms
 - Search/Sort algorithms
 - Graphs
 - Graph Algorithms (Traversal, MST, shortest-paths)
 - Time/Space Complexity
 - Programming Skills: Java/Python/MATLAB
 - E.g., Being able to crawl a website with Java
 - Basic knowledge of probability, statistics, calculus, and linear algebra
 - Expectation, variance, standard deviation,
 - Eigenvalue computation, determinants, characteristic equation
 - Basic differentiation, integration, and differential equations

背景知识

- 自然语言处理基础
- 信息检索基础
- 机器学习基础

自然语言处理

- 情感分析：
 - 文本情感分析（也称为意见挖掘）是指用自然语言处理、文本挖掘以及计算机语言学等方法来识别和提取原素材中的主观信息。
- 自动文摘：
 - 从相关文档中自动生成或者提取文字概括含义
- 文本分类：
 - 自动将文本根据预定义的label进行标记

自然语言处理

- 知识图谱
- 自然语言生成
- 问答&对话系统

信息检索基础

- 推荐系统：
 - 推荐系统是一种信息过滤系统，用于预测用户对物品的“评分”或“偏好”。
- 个性化检索与排序：个性化的排序结果，千人前面。。。。
- 用户画像：用户的年龄性别等

机器学习基础

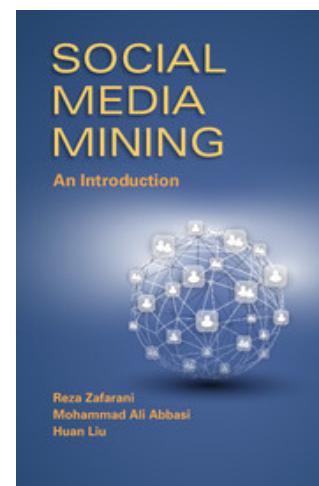
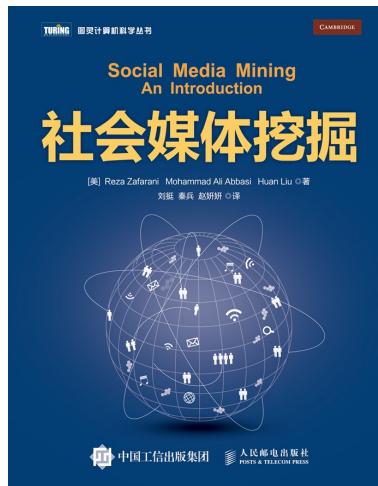
- 神经网络 Deep neural networks
- 主题建模 topic modeling
- 其他相关的工具：随机梯度下降，EM算法，近似算法，MCMC采样。。

Text and Reference Books

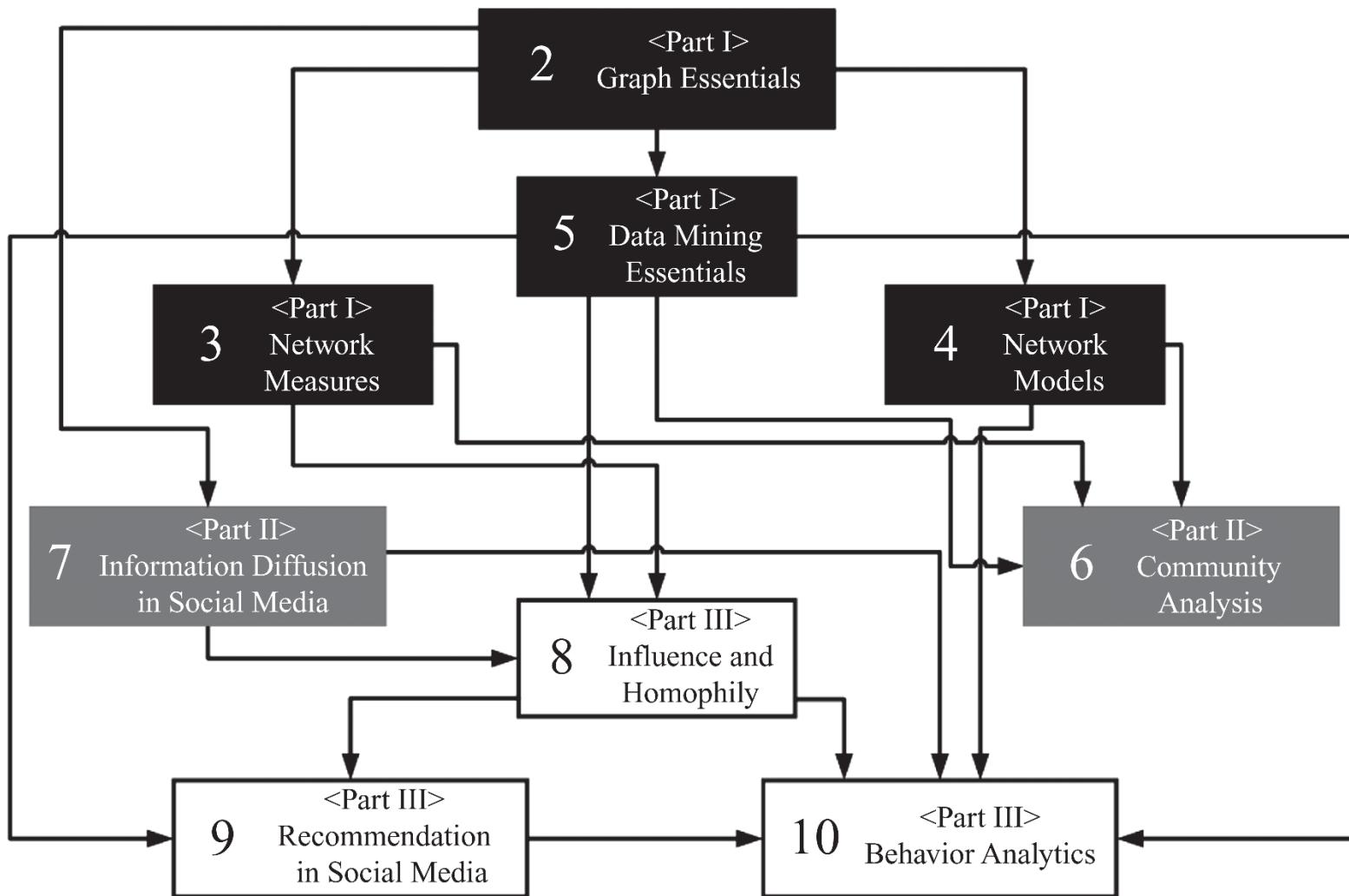
Primary Reference:

Social Media Mining, Reza Zafarani, Mohammad-Ali Abbasi, Huan Liu, Cambridge University Press 2014,

- Available at <http://socialmediamining.info>



Overview – Dependency Graph



课程内容

- 基础部分
 - 图结构与算法
 - 网络度量
 - 网络模型
 - 自然语言处理基础
- 应用部分
 - 社区发现
 - 社交网络中的信息传播
 - 社交媒体内容分析
 - 基于社交的推荐系统
 - 社交行为分析

评分系统

- Lectures
 - Experienced researchers or practitioners may be invited as guest instructors for specific topics.
- 课程作业 (20%) – 4次
 - Conceptual; deep thinking required
- 项目 (10%+10%)
 - Two projects
 - Group project + personal project
- 期末考试 (60%)

辅助教材和其他参考读物

- Monitoring social media: summarization, classification, and recommendation, Zhaochun Ren, University of Amsterdam
- Learning from Multiple Social Networks. Liqiang Nie, Xuemeng Song, Tat-Seng Chua. Synthesis Lectures on Information Concepts, Retrieval, and Services. Morgan & Claypool Publishers, 2016.
- Social media mining papers at SIGIR/KDD/WWW/WSDM/ACL/EMNLP, etc.

Outline

- 个人简介
- 课程简介
- 课程设置与安排
 - 课程目标
 - 背景知识
 - 数据与code

Short Break (10min)

- Lecture1: Introduction to Social Media Mining

Social Media

Definition

Social Media is the use of electronic and Internet tools for the purpose of sharing and discussing information and experiences with other human beings in more efficient ways.

Social Media Landscape 2015



FredCavazza.net

Social Media: Examples

- A wiki article
- Web reviews and ratings of a popular pizza place in your city
 - E.g., Yelp.com
- An online social network of your professional contacts
 - E.g., Facebook.com, LinkedIn.com
- An iPhone application that informs you where parking is likely available
 - FasPark

Types of Social Media

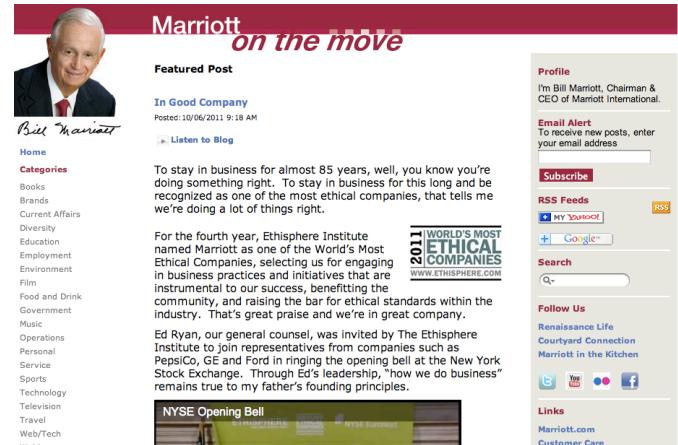
- Online Social Networking
- Publishing
 - Blogging
 - Wiki
- Micro blogging
- Social News
- Social Bookmarking
- Media Sharing
 - Video Sharing
 - Photo Sharing
 - Podcast Sharing
- Opinion, Review, and Ratings Websites
- Answers
- Entertainment



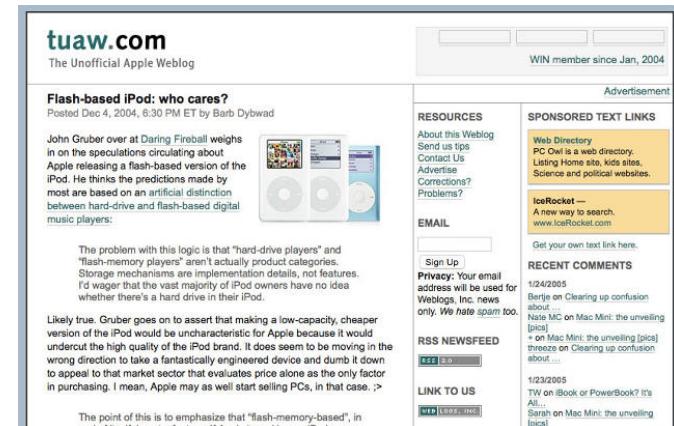
Blogging

A blog is a journal-like website for users, a.k.a. bloggers, to contribute textual and multimedia content, arranged in reverse chronological order

- Maintained both individually or by a community
 - See a tutorial at KDD
http://videolectures.net/kdd08_liu_briat/
- Usages:
 - Sharing information and opinions with friends and strangers
 - Disseminating subject-specific content
 - Who is the influential
http://videolectures.net/wsdm08_agarwal_ib/



The screenshot shows the Marriott International website. At the top, there's a banner with a photo of Bill Marriott and the text "Marriott on the move". Below the banner, there's a "Featured Post" section with a photo of Bill Marriott and the title "In Good Company". The post was posted on 10/06/2011 at 9:18 AM and has a link to "Listen to Blog". A sidebar on the right lists "Categories" such as Books, Brands, Current Affairs, Diversity, Education, Employment, Environment, Film, Food and Drink, Government, Music, Operations, Personal, Service, Sports, Technology, Television, Travel, Web/Tech, and more. Another sidebar on the far right contains links for "Profile", "Email Alert", "RSS Feeds" (with links to My Yahoo!, Google, and RSS), "Search", "Follow Us" (with links to Renaissance Life, Courtyard Connection, Marriott in the Kitchen, and social media icons for Twitter, YouTube, LinkedIn, and Facebook), and "Links" (with links to Marriott.com and Customer Care). A banner for "WORLD'S MOST ETHICAL COMPANIES" is also visible.



The screenshot shows the tuaw.com website, which is described as "The Unofficial Apple Weblog". The main article is titled "Flash-based iPod: who cares?" and was posted on Dec 4, 2004, at 6:30 PM ET by Barb Dybwad. The article discusses John Gruber's opinion on the potential impact of Apple releasing a flash-based version of the iPod. It includes a photo of two iPods. To the right of the article, there's a sidebar with "RESOURCES" (links to About this Weblog, Send us tips, Contact Us, Advertising, Corrections, and Problems) and "SPONSORED TEXT LINKS" (links to Web Directory, PC Our, Listing Home site, kids sites, Science and political websites, and IceRocket). There are also sections for "EMAIL" (Sign Up, Privacy), "RSS NEWSFEED" (links to RSS 2.0 and Atom), and "LINK TO US" (links to RSS 2.0 and Atom). A "RECENT COMMENTS" sidebar shows recent posts from users like Berlie, Nate MC, and Sarah on Mac Mini. The footer of the page includes copyright information and a link to "Get your own text link here".

Microblogging

Microblogging can be considered as a counterpart to blogging, but with limited content

- Usage
 - communication medium
 - social interaction
 - citizen journalism
- Service Providers:
 - Twitter
 - Google buzz



Wiki

A wiki is a collaborative editing environment that allows users to develop Web pages using a simplified markup language

- Wikipedia allows interested individuals to collaboratively develop articles on a variety of subjects.
- Using the wisdom of crowds effectively, it has become a comprehensive repository of information useful to a variety of individuals

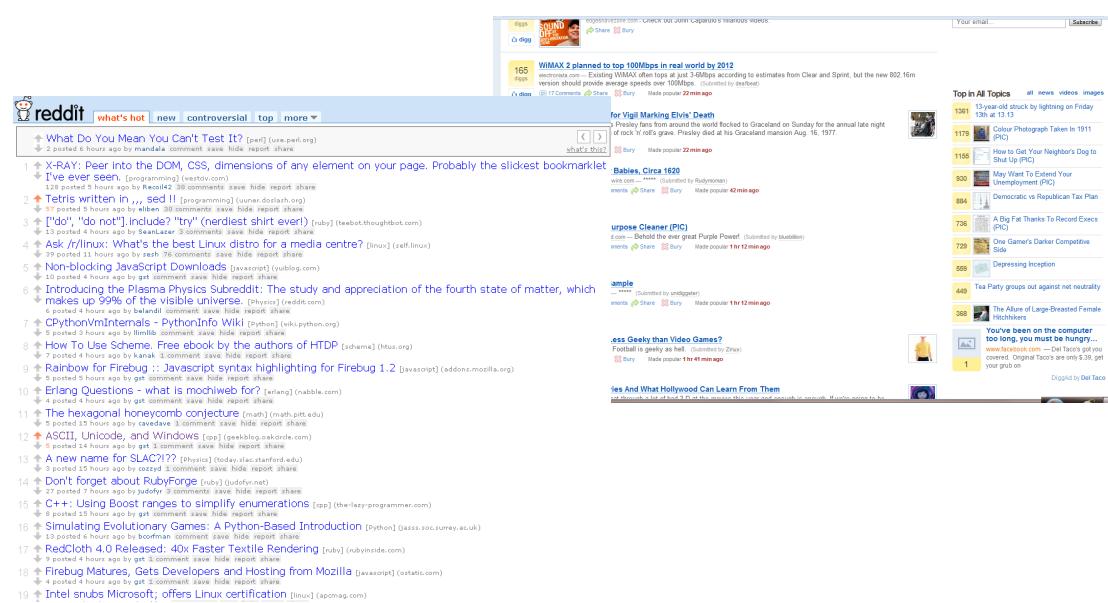
The screenshot shows the Wikipedia article for "Linux". The page title is "Linux" and the subtitle is "The Free Encyclopedia". The main content discusses the Linux kernel and its history. A sidebar on the left contains links like "Main page", "Contents", and "Recent changes". Another sidebar on the right provides details about the Linux kernel, including its latest stable release (Kernel 2.6.32.1), supported platforms (x86, MPPC, ppc64, SPARC, DEI Alpha, Itanium, PowerPC, ARM, m68k, PA-RISC, x390, SuperH, M32R and more), and its kernel type (Monolithic kernel). The footer includes a link to the Linux kernel's GitHub repository.

Social News

Social News refers to the sharing and selection of news stories and articles by a community of users.

- Users can share articles that they believe would interest the community
- Samples:

- Digg.com
- Slashdot
- Fark
- Reddit



Social Bookmarking

Social Bookmarking sites allow users to bookmark web content for storage, organization and sharing.

- These bookmarks can be tagged with metadata to categorize and provide context to the shared content, allowing users to organize information making it easy to search and identify relevant information.
- Samples
 - Delicious.com
 - StumbleUpon.com

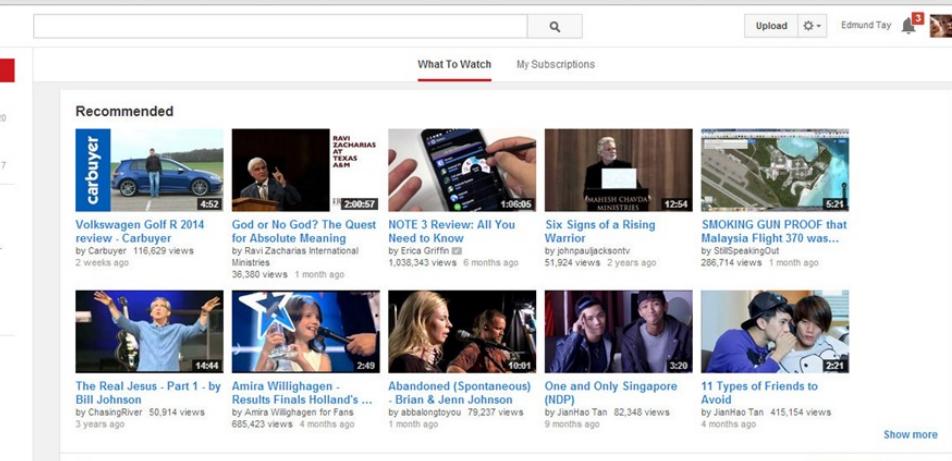
The screenshot shows the Delicious.com homepage. At the top, there's a navigation bar with links for "It's Free!", "Join Now", and "Sign In". Below the header, a large banner features the text "The tastiest bookmarks on the web. Save your own or see what's fresh now!" with a "Learn More" button and three small cupcake icons. A search bar at the bottom of the banner says "Search the biggest collection of bookmarks in the universe...". To the right of the search bar are buttons for "Search Delicious" and "Search". Below the banner, tabs for "Popular Bookmarks" and "Explore Tags" are visible. The main content area displays a list of popular bookmarks with their titles, save counts, and tags. To the right of the list is a sidebar titled "Popular Tags" with a scrollable list of tags like "design", "blog", "video", etc. At the bottom of the page, there's a footer with links for "delicious", "about", "blog", "terms of service", "privacy policy", "copyright policy", "forums", "support", and "What's new?".

Media Sharing

Media sharing is an umbrella term that refers to the sharing of a variety of media on the web.

Users share such multimedia content of possible interest to others

- Samples:
 - Video Sharing:
 - YouTube.com
 - Photo Sharing:
 - Flickr.com, picasa.com
 - Document Sharing:
 - Scribd.com,
Slideshare.com
 - Livecasting:
 - Justin.tv, Ustream.com



Opinion, Review, and Ratings Websites

Opinion, review, and ratings websites are websites whose primary function is to collect and publish user-submitted content in the form of subjective commentary on existing products, services, entertainment, businesses, places, etc. Some commercial sites may serve a secondary purpose as review sites by publishing product reviews submitted by customers.

- Examples

- Cnet.com
- Epinions.com
- yelp.com
- tripadvisor.com

The screenshot shows a Yelp search results page for "sf, ca". The search bar has "sf, ca" entered. Below the search bar, there are tabs for "Welcome", "About Me", "Write a Review", "Find Friends", "Messaging", "Talk", and "Events". On the right, there is a "Member Search" button. The main content area shows a product listing for a "Croissant" at "Tartine Bakery". The listing includes a price of "\$3.85", three small images of the croissant, and a link to "See more photos". Below the image, there is a section for reviews. A review by "Stephanie S." from "Loma Linda, CA" is shown, dated "10/23/2012", with a rating of 5 stars. The review text reads: "This was our first stop from the airport and we were starving! The line was long, but it went pretty fast. This was our first time here and we couldn't decide what to order. We tried the morning bun, chocolate and almond croissant, bread pudding, & the chocolate eclair. Everything was delicious, but the morning bun was sooo amazing. I loved the hints of citrus and the flakiness of the bun. I made my hubby go back & buy me another one to save for later. Oh, Tartine! I wish you were also located in So. Cal." Below the review, there are buttons for "Write a Review", "Add a photo", "Complement", "Send Message", and "Follow This Reviewer". At the bottom of the review section, there are links for "Was this review...?", "Useful", "Funny", "Cool", "Bookmark", "Send to a Friend", "Link to This Review", and "Flag this review". To the right of the main listing, there is a sidebar titled "Menu for Tartine Bakery" under "Breakfast Pastries". It lists several items with their names, reviews, photos, and prices:

Item	Reviews	Photos	Price
Croissant	1147 reviews	49 photos	\$3.85 Popular
Frangipane Croissant	63 reviews	5 photos	\$4.50
Double Pain Au Chocolat	78 reviews	12 photos	\$4.50
Morning Buns	355 reviews	13 photos	\$3.85
Buttermilk Scones	3 reviews	1 photo	\$3.25
Tea Cake	36 reviews	9 photos	\$3.75
Bread Pudding	898 reviews	85 photos	Price details Popular
Pain Au Jambon	24 reviews	8 photos	\$4.95
Gougere	130 reviews	14 photos	\$3.50
Cake Aux Olives	9 reviews		\$4.95
Quiche	379 reviews	29 photos	Price details Popular
Muesli			Price details

Socially-Provided Answers

In these sites, users who require certain guidance, advice or knowledge can ask questions. Other users from the community can answer these questions based on knowledge acquired from previous experiences, personal opinions or from relevant research.

- Unlike review and opinion sites, which contain self-motivated contribution of opinions, answer sites contain knowledge shared in response to a specific query.
- Samples:
 - WikiAnswers, Yahoo Answers, Quora

 Add Question

Question added to topic Google Analytics:

[What percentage of visits would Omniture / Google Analytics / Coremetrics etc miss?](#)

Assuming client-side integration, compared with the numbers from the web servers and proxy logs.

Follow · Repost · 0 Answers · 5:55pm

Answer added in topic Google Analytics:

[How can I track Pinterest in Google Analytics?](#)

1 Ross Allen, Front End Engineer at Airbnb

Their Javascript pinit.js file (<http://assets.pinterest.com/js/pinit.js>) doesn't seem to add any callbacks, so the best you can do is track clicks on the 'Pin It' button in Goo... [\(more\)](#)

Upvote · Repost · 2 Answers · 5:17pm

Answer added in topic Google Analytics:

[Google Analytics: Why would someone from an email marketing company tell me that Google analytics does not track visits from Mac users?](#)

2 Anon User

The person was seeing if you were gullible enough to be a good fit with their product.

Sales 101.

Upvote · Repost · 4 Answers · 3:52pm

[Share Topic](#) • [Invite People](#)

[Twitter](#) [Facebook](#) [Quora](#)

Top Answerers

 **Mike Sullivan**
20 Answers

 **Ozberk Olcer**
20 Answers Director of Web Analytics in SEM AS. (Google Analytics Certified Partner)

 **Shay Sharon**
22 Answers

 **AJ Kohn**
17 Answers

 **Christopher O'Donnell**
11 Answers

Followed by 5455 People



Main Characteristics

- **Participation**

- social media encourages contributions and feedback from everyone who is interested. It blurs the line between media and audience.

- **Openness**

- most social media services are open to feedback and participation. They encourage voting, comments and the sharing of information. There are rarely any barriers to accessing and making use of content – password-protected content is frowned on.

- **Conversation**

- whereas traditional media is about “broadcast” (content transmitted or distributed to an audience) social media is better seen as a two-way conversation.

- **Community**

- social media allows communities to form quickly and communicate effectively. Communities share common interests, such as a love of photography, a political issue or a favorite TV show.

- **Connectedness**

- Most kinds of social media thrive on their connectedness, making use of links to other sites, resources and people.

Social Media Mining is the process of representing, analyzing, and extracting meaningful patterns from social media data

Social Media Mining Challenges

1. Big Data Paradox

1. Social media data is big, yet not evenly distributed.
2. Often little data is available for an individual

2. Obtaining Sufficient Samples

1. Are our samples reliable representatives of the full data?

3. Noise Removal Fallacy

1. Too much removal makes data more sparse
2. Noise definition is relative and complicated and is task-dependent

4. Evaluation Dilemma

1. When there is no ground truth, how can you evaluate?

Social Networks

Online Social Networking

Online Social Networks are web-based services that allow individuals and communities to connect with real world friends and acquaintances online

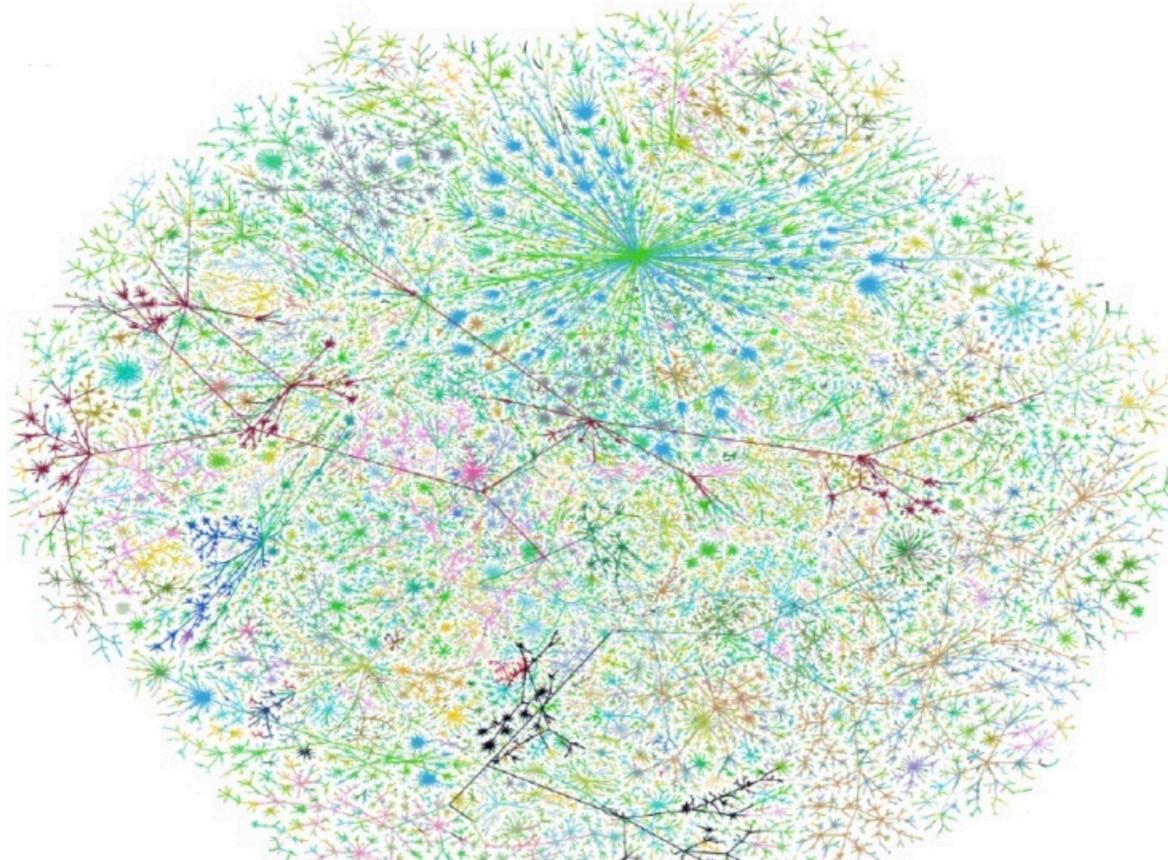
- Interactions
 - Friendship interaction
 - Friends, like, comments, ...
 - Media Sharing
 - Sending and receiving messages

• Examples

- Facebook.com
- MySpace.com
- Bebo.com
- Orkut.com

The image displays three separate screenshots of online social networking platforms:

- MySpace Profile (Pei Pei):** This screenshot shows a MySpace profile for a user named "Pei Pei". The profile includes basic information like location (Seattle, United States) and profile views (208). It features a large photo of a person's face, a music player showing "Greenville" by Pei Pei, and a blog entry titled "I Don't Need To Be Rich" by Pei Pei. The profile also lists upcoming shows at "Freshly's Cafe" on March 20, 2010.
- Facebook Page (Barack Obama):** This screenshot shows the official Facebook page for Barack Obama. It features a large profile picture of him, a post about voting, and a sidebar with news items and links related to his campaign.
- MySpace Profile (Pei Pei):** This screenshot shows another MySpace profile for "Pei Pei", identical to the first one but with a different URL (www.myspace.com/peixx). It includes a bio mentioning influences like "Too many to list them all", "The White Stripes", "U2", "Zee Avi", "A Fine Frenzy", "Black Whale", "Damien Rice", and "more.", and a "Type of Label" section indicating "Unsigned".



**We will never be able to model and
predict these systems unless we
understand the networks behind them!**

社交网络的影响力



or
+
by

 Google

Cisco

Facebook

Amazon

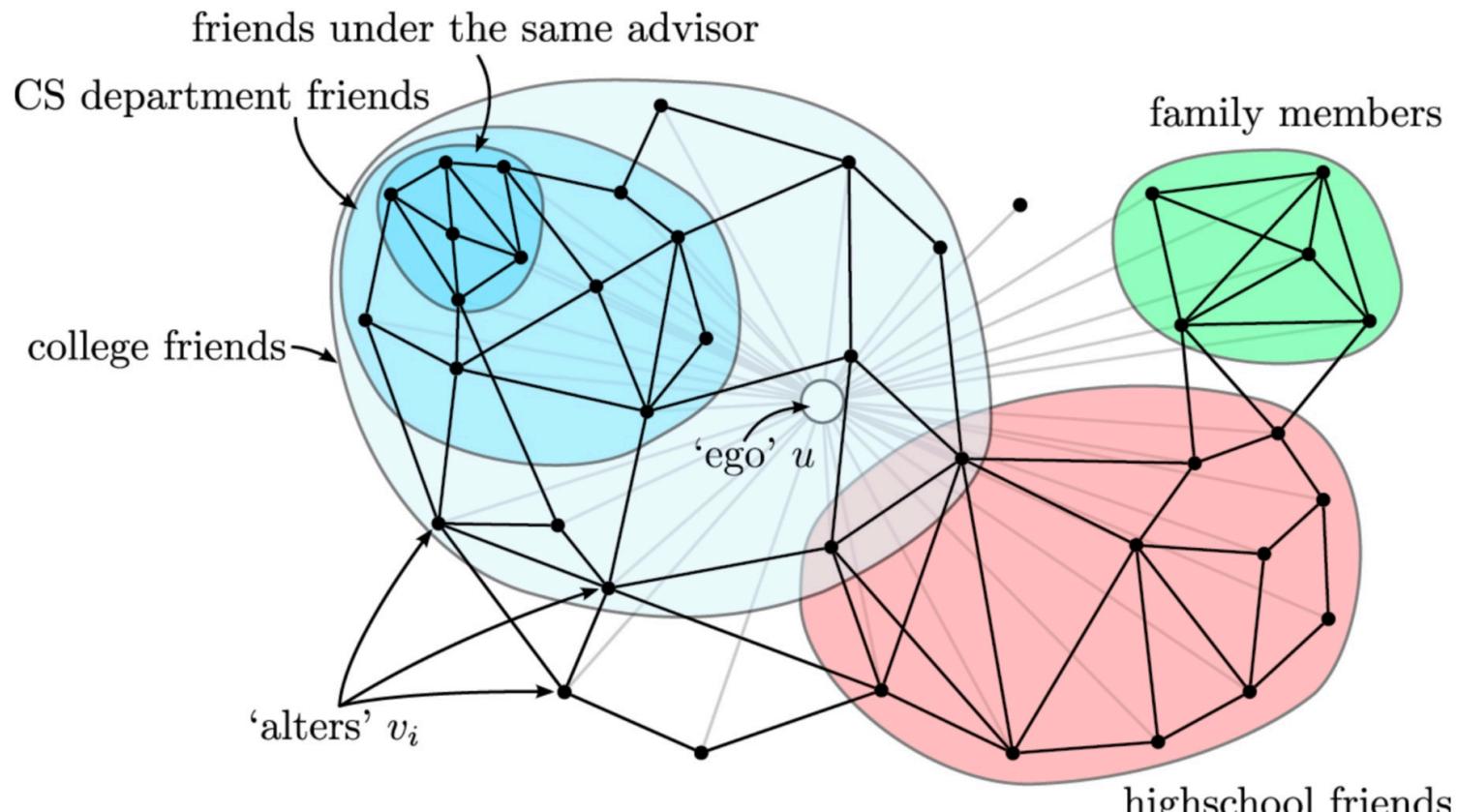
Social Networks



Facebook social graph

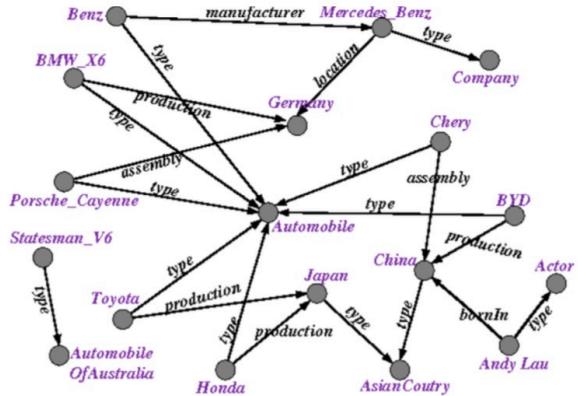
4-degrees of separation [Backstrom-Boldi-Rosa-Ugander-Vigna, 2011]

Task1: Social Circle Detection

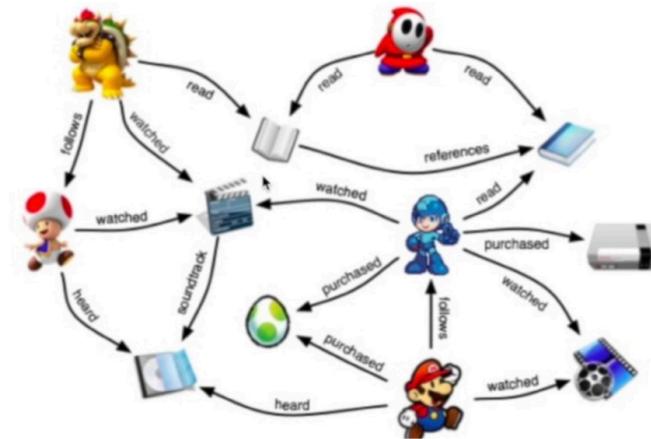


Discover circles and why they exist

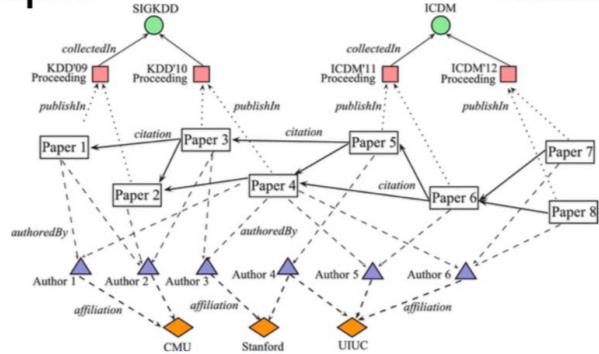
Task2: Information&Knowledge discovery



Knowledge Graphs



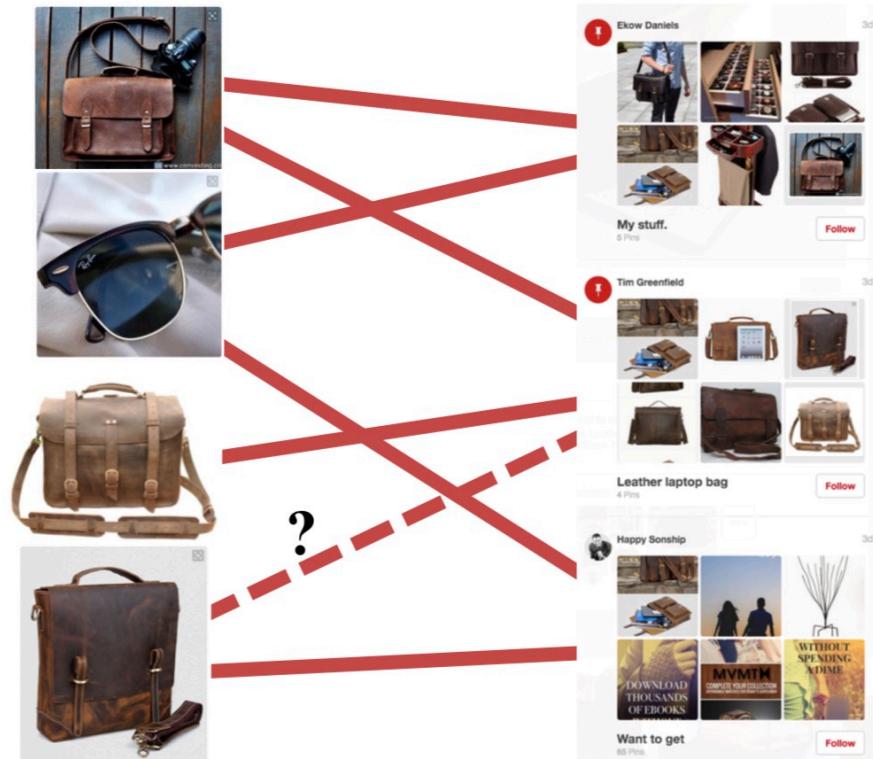
Heterogeneous Graphs



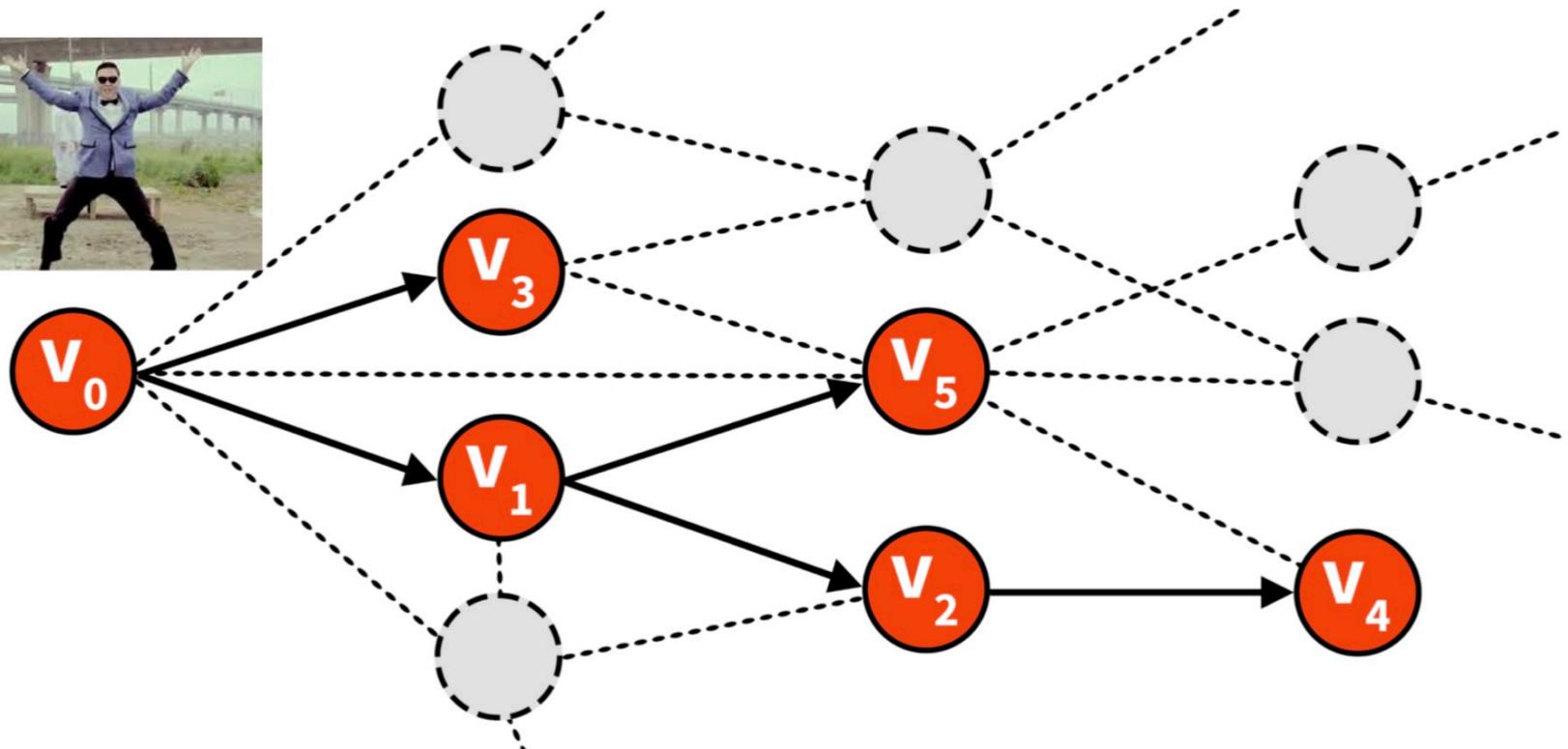
Multimodal Graphs

Task3: Link Prediction

Content
recommendation is
link prediction



Task4: predicting virality



Information cascade in social networks

分析社交网络的方法

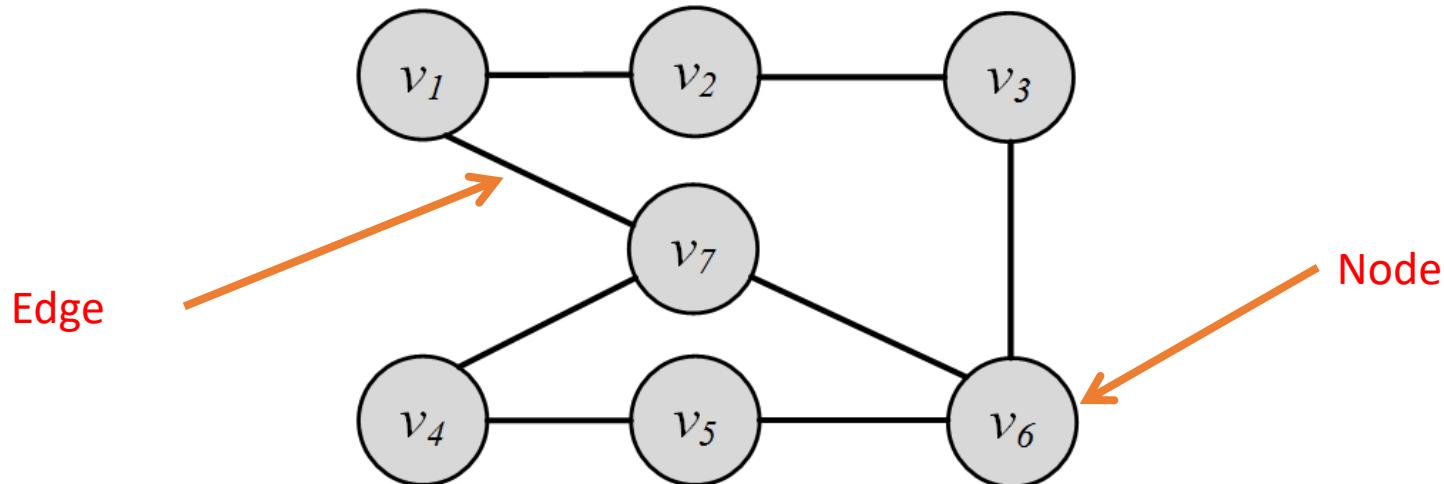
- **Predict the type/color of a given node**
 - Node classification
- **Predict whether two nodes are linked**
 - Link prediction
- **Identify densely linked clusters of nodes**
 - Community detection
- **Measure similarity of two nodes/networks**
 - Network similarity

Graph Basics

Nodes and Edges

A network is a graph, or a collection of points connected by lines

- Points are referred to as **nodes**, **actors**, or **vertices** (plural of **vertex**)
- Connections are referred to as **edges** or **ties**



Nodes or Actors

- In a friendship social graph, nodes are people and any pair of people connected denotes the friendship between them
- Depending on the context, these nodes are called nodes, or actors
 - In a web graph, “*nodes*” represent sites and the connection between nodes indicates web-links between them
 - In a social setting, these nodes are called actors

• The size of the graph is $V = \{v_1, v_2, \dots, v_n\}$

$$|V| = n$$

Edges

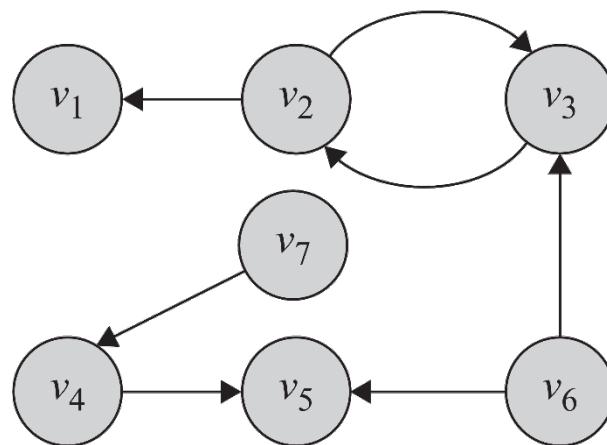
- Edges connect nodes and are also known as **ties** or **relationships**
- In a social setting, where nodes represent social entities such as people, edges indicate internode relationships and are therefore known as relationships or (social) ties

$$E = \{e_1, e_2, \dots, e_m\}$$

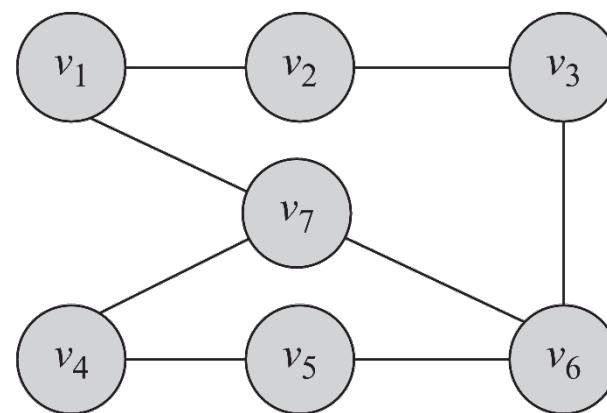
- Number of edges (size of the edge-set) is denoted as $|E| = m$

Directed Edges and Directed Graphs

- Edges can have directions. A directed edge is sometimes called an **arc**



(a) Directed Graph



(b) Undirected Graph

- Edges are represented using their end-points $e(v_2, v_1)$.
- In undirected graphs both representations are the same

Neighborhood and Degree (In-degree, out-degree)

For any node v , in an undirected graph, the set of nodes it is connected to via an edge is called its neighborhood and is represented as $N(v)$

- In directed graphs we have incoming neighbors $N_{in}(v)$ (nodes that connect to v) and outgoing neighbors $N_{out}(v)$.

The number of edges connected to one node is the degree of that node (the size of its neighborhood)

- Degree of a node i is usually presented using notation d_i

In Directed graphs:

d_i^{in} • In-degrees is the number of edges pointing towards a node

d_i^{out} • Out-degree is the number of edges pointing away from a node

Degree and Degree Distribution

- **Theorem 1.** The summation of degrees in an undirected graph is twice the number of edges

$$\sum d_i = 2|E|$$

- **Lemma 1.** The numberⁱ of nodes with odd degree is even
- **Lemma 2.** In any directed graph, the summation of in-degrees is equal to the summation of out-degrees,

$$\sum_i d_i^{out} = \sum_j d_j^{in}$$

Degree Distribution

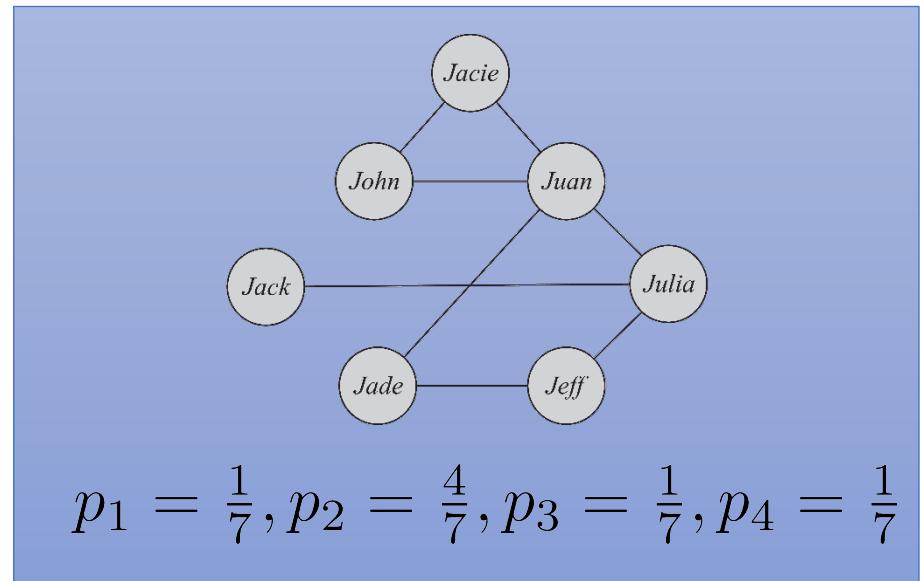
When dealing with very large graphs, how nodes' degrees are distributed is an important concept to analyze and is called ***Degree Distribution***

$$\pi(d) = \{d_1, d_2, \dots, d_n\} \quad (\text{Degree sequence})$$

$$p_d = \frac{n_d}{n}$$

n_d is the number of nodes with degree d

$$\sum_{d=0}^{\infty} p_d = 1$$



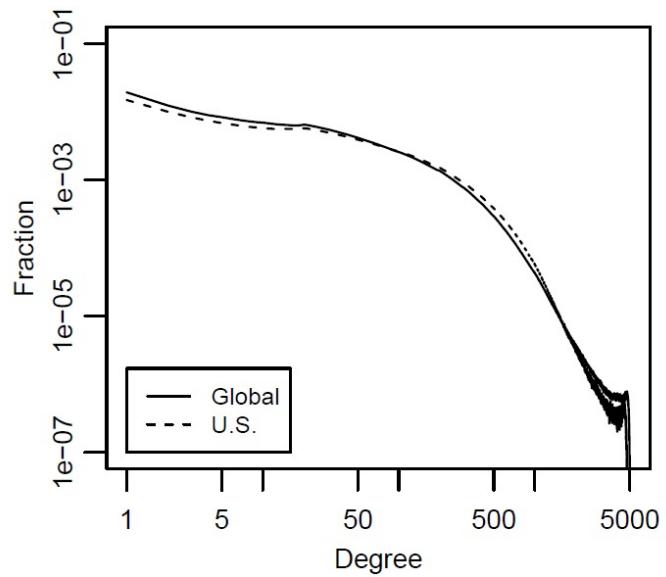
Degree Distribution Plot

The x -axis represents the degree and the y -axis represents the fraction of nodes having that degree

- On social networking sites

There exist many users with few connections and there exist a handful of users with very large numbers of friends.

(Power-law degree distribution)



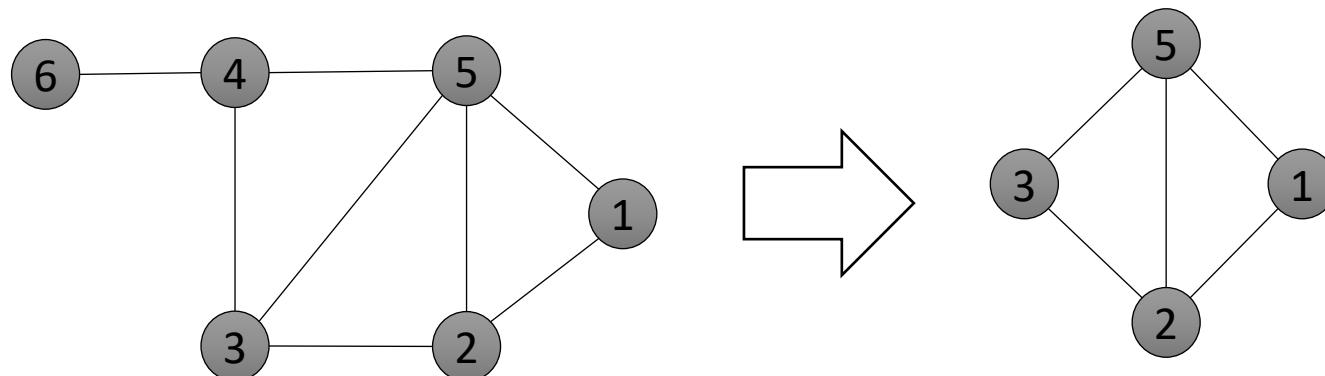
**Facebook
Degree Distribution**

Subgraph

- Graph G can be represented as a pair $G(V, E)$ where V is the node set and E is the edge set
- $G'(V', E')$ is a subgraph of $G(V, E)$

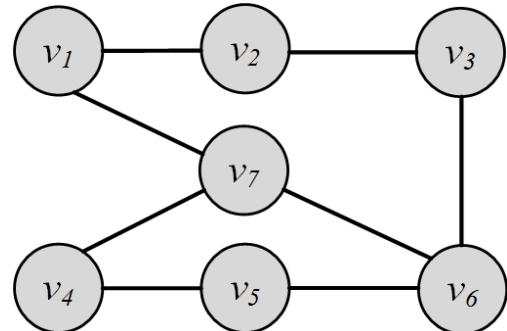
$$V' \subseteq V$$

$$E' \subseteq (V' \times V') \cap E$$



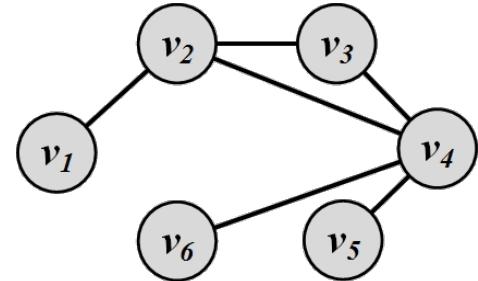
Graph Representation

- Adjacency Matrix
- Adjacency List
- Edge List



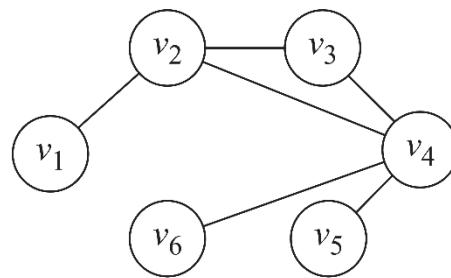
Graph Representation

- Graph representation is straightforward and intuitive, but it cannot be effectively manipulated using mathematical and computational tools
- We are seeking representations that can store these two sets in a way such that
 - Does not lose information
 - Can be manipulated easily by computers
 - Can have mathematical methods applied easily



Adjacency Matrix (a.k.a. sociomatrix)

$$A_{ij} = \begin{cases} 1, & \text{if there is an edge between nodes } v_i \text{ and } v_j \\ 0, & \text{otherwise} \end{cases}$$



(a) Graph

	v ₁	v ₂	v ₃	v ₄	v ₅	v ₆
v ₁	0	1	0	0	0	0
v ₂	1	0	1	1	0	0
v ₃	0	1	0	1	0	0
v ₄	0	1	1	0	1	1
v ₅	0	0	0	1	0	0
v ₆	0	0	0	1	0	0

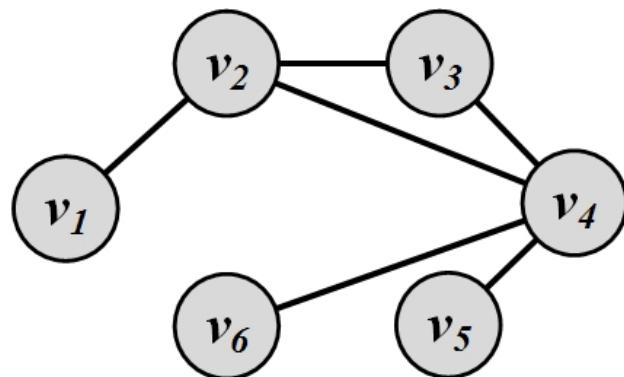
(b) Adjacency Matrix

Diagonal Entries are self-links or loops

Social media networks have
very **sparse** Adjacency matrices

Adjacency List

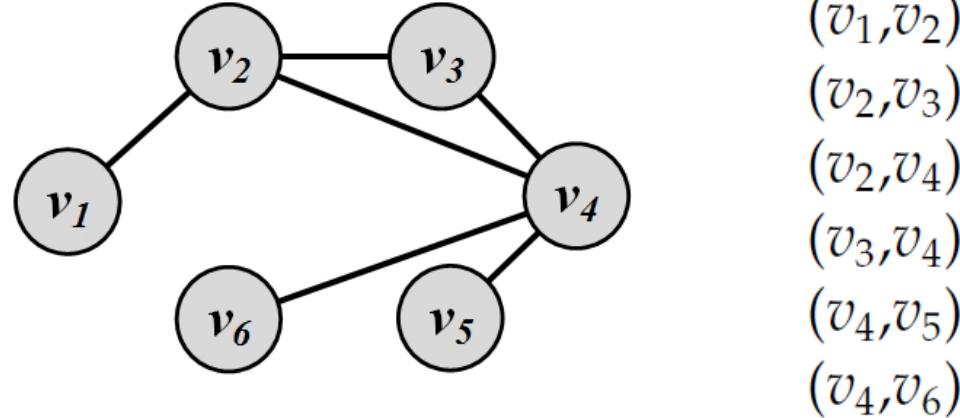
- In an adjacency list for every node, we maintain a list of all the nodes that it is connected to
- The list is usually sorted based on the node order or other preferences



Node	Connected To
v_1	v_2
v_2	v_1, v_3, v_4
v_3	v_2, v_4
v_4	v_2, v_3, v_5, v_6
v_5	v_4
v_6	v_4

Edge List

- In this representation, each element is an edge and is usually represented as (u, v) , denoting that node u is connected to node v via an edge



Types of Graphs

- Null, Empty,
Directed/Undirected/Mixed,
Simple/Multigraph, Weighted, Signed
Graph, Webgraph

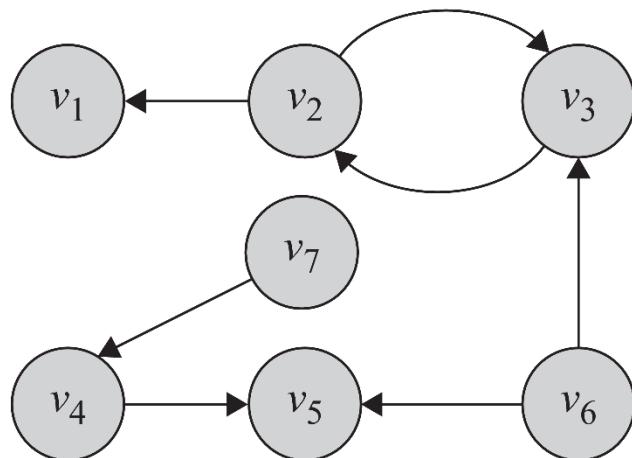
Null Graph and Empty Graph

- A **null graph** is one where the node set is empty (there are no nodes)
 - Since there are no nodes, there are also no edges

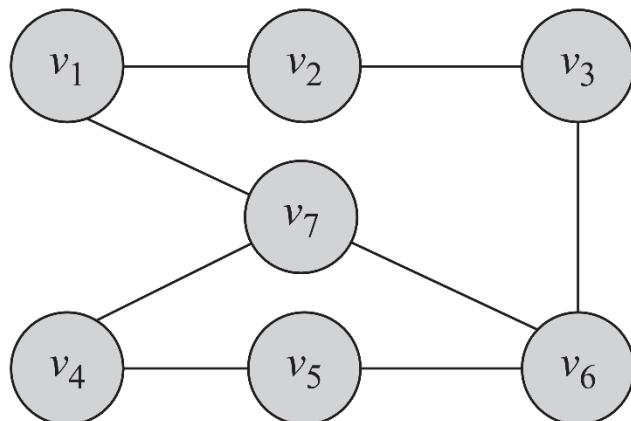
$$G(V, E), V = E = \emptyset$$

- An **empty graph** or **edge-less graph** is one where the edge set is empty,
 $E = \emptyset$
- The node set can be non-empty.
 - A null-graph is an empty graph.

Directed/Undirected/Mixed Graphs



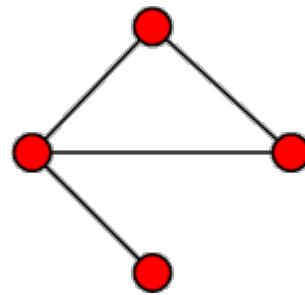
- The adjacency matrix for directed graphs is often not symmetric ($A \neq A^T$)
 - $A_{ij} \neq A_{ji}$
 - We can have equality though



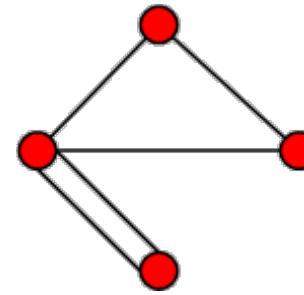
The adjacency matrix for undirected graphs is symmetric ($A = A^T$)

Simple Graphs and Multigraphs

- Simple graphs are graphs where only a single edge can be between any pair of nodes
- Multigraphs are graphs where you can have multiple edges between two nodes and loops



Simple graph

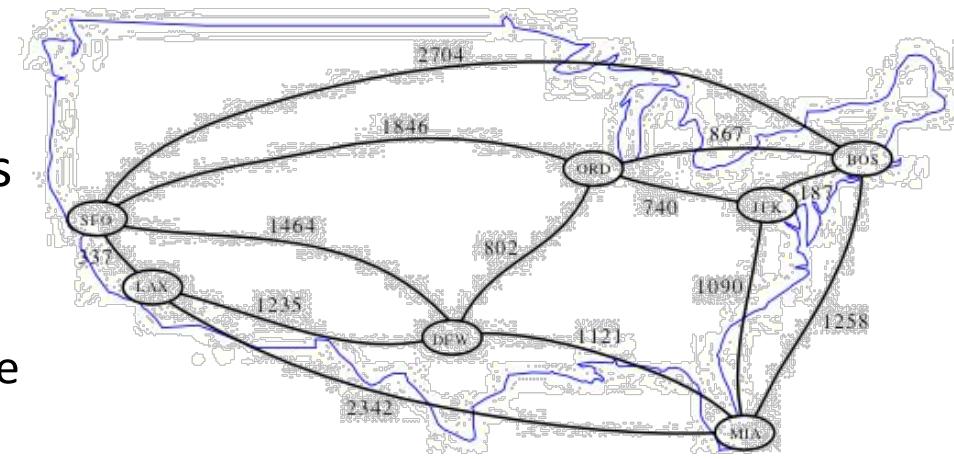


Multigraph

- The adjacency matrix for multigraphs can include numbers larger than one, indicating multiple edges between nodes

Weighted Graph

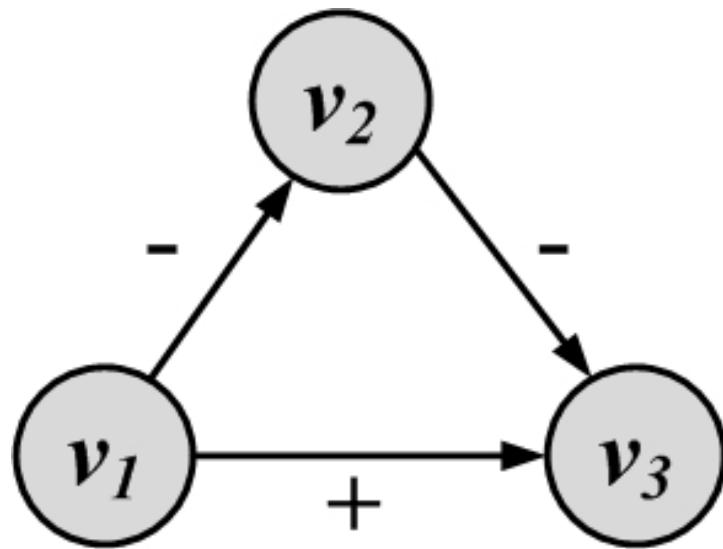
- A weighted graph $G(V, E, W)$ is one where edges are associated with weights
 - For example, a graph could represent a map where nodes are airports and edges are routes between them
 - The weight associated with each edge could represent the distance between the corresponding cities



$$A_{ij} = \begin{cases} w_{ij} \text{ or } w(i, j), w \in R \\ 0, \text{ There is no edge between } v_i \text{ and } v_j \end{cases}$$

Signed Graph

- When weights are binary (0/1, -1/1, +/-) we have a **signed** graph



- It is used to represent **friends** or **foes**
- It is also used to represent **social status**

Webgraph

- A webgraph is a way of representing how internet sites are connected on the web
- In general, a web graph is a directed multigraph
- Nodes represent sites and edges represent links between sites.
- Two sites can have multiple links pointing to each other and can have loops (links pointing to themselves)

总结

- 社交媒体简介
- 社交网络简介
- 图基础
- Next...
 - 图结构
 - 自然语言处理基础