

# Рекомендательные системы и сети

DS Minor

Алена Владимировна Суворова

# Рекомендательные системы

алгоритмы, которые пытаются предсказать, какие объекты (фильмы, музыка, книги, новости, веб-сайты) будут интересны пользователю, имея определенную информацию о его профиле.

- Amazon
- Last.fm, Pandora
- *Google, Yahoo!, Яндекс*

# Два способа рекомендовать

**Ориентация на  
пользователя**



*Пользователи, похожие  
на вас, также  
слушают...*

**Ориентация на  
продукт**



*С этим товаром также  
покупают...*

# How to solve a data analysis problem



Take a problem



Identify Data sources & Acquire data



Data Cleansing



Model Validation



Build Statistical Model



Exploratory Analysis



Visualize results & Do predictions

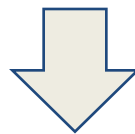
# Основные стратегии создания

**коллаборативная  
фильтрация**



информация  
о поведении пользователей  
в прошлом, выбираются  
пользователи с похожим  
поведением

**фильтрация на основе  
содержания**



профили пользователей и  
объектов, используются  
похожие профили

# Алгоритмы: коллаборативная фильтрация

Используются известные предпочтения (оценки) группы пользователей для прогнозирования неизвестных предпочтений другого пользователя

- вычислить похожесть всех пар объектов
- для каждого объекта найти k наиболее похожих
- для каждого пользователя определить объекты, наиболее близкие к тем, которые он оценил

# Сети и рекомендательные системы

Использование данных из сетей для рекомендаций

Рекомендации в сетях («вы можете знать»)

Социальные сети как платформа для представления рекомендаций

# Применение рекомендательных систем

Amazon

Кинопоиск

Интернет-магазины

...

Рекомендательные системы вокруг вас?



## Эхо-камеры

«Эхо-камера» -- это ситуация, в которой информация, идеи или убеждения отражаются и усиливаются передачей и повторением внутри «закрытой» системы, где альтернативные точки зрения цензурируются, запрещаются или слабо представлены по другим причинам

# Пузырь фильтров



Пузырь фильтров -- ситуация, в которой алгоритмы вебсайтов отображают вам только те новости и посты, которые, по их предсказанию, могут вам понравиться

# Персонализированный поиск Google

## Windows Phone (United States)

[www.windowsphone.com/](http://www.windowsphone.com/)

"I love my **Windows Phone**. The stuff that's important to me is pinned to my Start screen, like the Live Tile for my All Recipes app which updates regularly with ...

My Phone - Windows Phone 7 - Canada - India

## Windows Phone Apps+Games Store (United ... - WindowsPhone.com

[www.windowsphone.com/en-us/store](http://www.windowsphone.com/en-us/store)

Browse Apps+Games for Windows Phone. ... Popular and Famous. This collection features the most famous and innovative apps in the Store. First Row ...

## Windows Phone - Wikipedia, the free encyclopedia

[en.wikipedia.org/wiki/Windows\\_Phone](http://en.wikipedia.org/wiki/Windows_Phone)

**Windows Phone** is a family of mobile operating systems developed by Microsoft, and is the successor to its Windows Mobile platform, although incompatible with ...

## Phones | Windows Phone (United States)

[www.windowsphone.com/en-us/phones](http://www.windowsphone.com/en-us/phones)

Features - Phones - Apps+Games | News - How-to, Explore, My Phone. Have a Windows Phone? Sign in - My Phone - Find My Phone - Purchase History ...

## Windows Phone 8 - New Windows 8 Phones, Windows 7 ... - AT&T

[www.att.com](http://www.att.com) - Wireless - Devices

AT&T has the newest **Windows Phone 8** - Windows 7 Phones, including smartphones and devices from HTC, Samsung, and Nokia. Choose the best Windows 8 ...

## Windows Phone (@windowsphone) on Twitter

<https://twitter.com/windowsphone>

The Official **Windows Phone** Twitter Channel - keeping you updated with the latest ... "If iOS bores you and Android intimidates you, **WindowsPhone** will likely ...

## Windows Phone (United States)

[www.windowsphone.com/](http://www.windowsphone.com/)

"I love my **Windows Phone**. The stuff that's important to me is pinned to my Start screen, like the Live Tile for my All Recipes app which updates regularly with ...

My Phone - Windows Phone 7 - Canada - India

## Windows Phone Apps+Games Store (United ... - WindowsPhone.com


[www.windowsphone.com/en-us/store](http://www.windowsphone.com/en-us/store)

Browse Apps+Games for Windows Phone. ... Popular and Famous. This collection features the most famous and innovative apps in the Store. First Row ...

You've visited this page 2 times. Last visit: 10/29/12

## Yes, you can use Google+ through Internet Explorer on Windows ...

<https://plus.google.com/.../posts/5C5FmFeedyP>

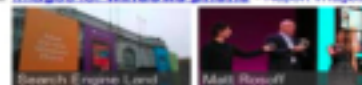
 Danny Sullivan - 46 minutes ago - Yes, you can use Google+ through Internet Explorer on **Windows Phone 8**, but it's not a pleasant experience. I've found. I can haz app please

## Windows Phone - Wikipedia, the free encyclopedia

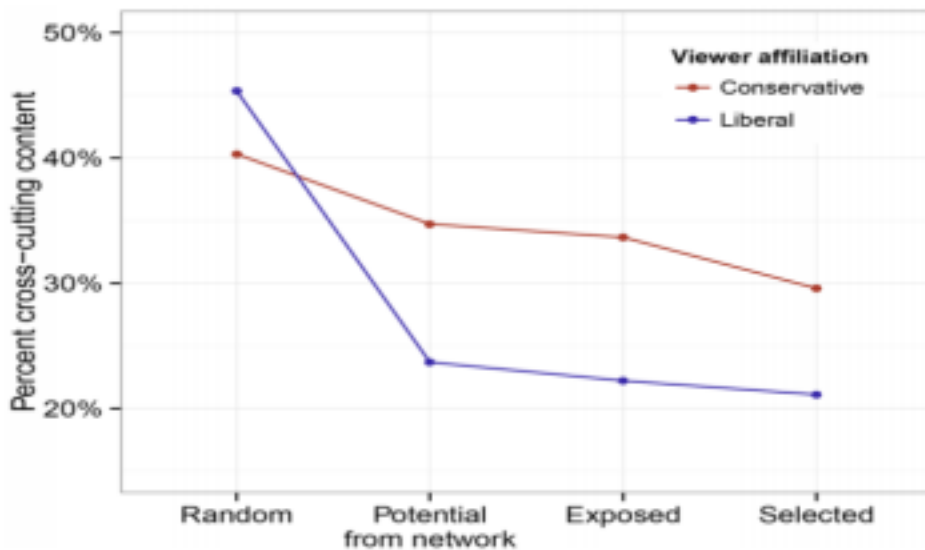
[en.wikipedia.org/wiki/Windows\\_Phone](http://en.wikipedia.org/wiki/Windows_Phone)

**Windows Phone** is a family of mobile operating systems developed by Microsoft, and is the successor to its Windows Mobile platform, although incompatible with ...

## Images for windows phone - Report Images



## Алгоритм ранжирования новостей Facebook



**Сильные и слабые связи**

# Цели изучения

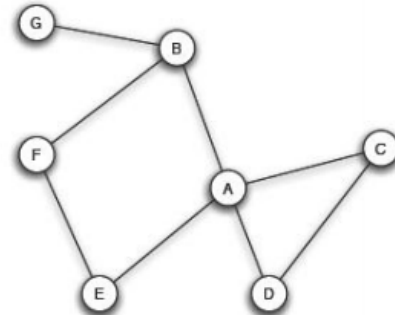
- понять, как информация распространяется в сети
- какие роли есть у вершин сети
- как структурные свойства сети влияют на изменение сети во времени

# «Поиск работы»

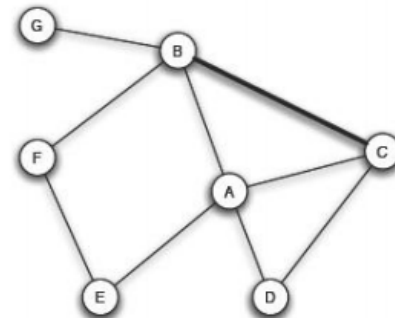
- Марк Грановеттер (1960е)
  - опрос тех, кто недавно поменял работу, о том, как они узнали о вакансии
  - во многих случаях — через личные контакты
  - причем часто такие «полезные» знакомства — это просто знакомые (слабые связи), а не близкие друзья (сильные связи)
- Хотя
  - друзья более мотивированы помочь
  - почему же благодарить нужно не их?

# Triadic Closure

- Если у вас есть два друга, то большие шансы, что они подружатся между собой



(a) Before B-C edge forms.



(b) After B-C edge forms.

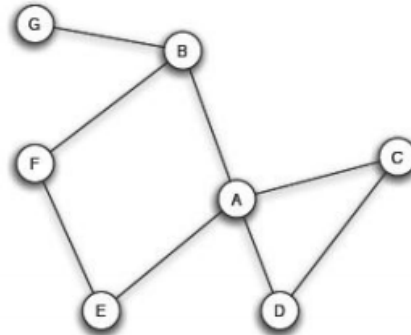


# Почему образуется связь?

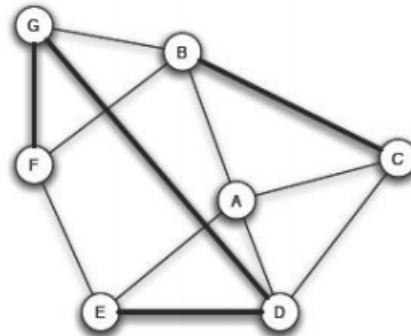
- больше возможностей встретиться
- больше доверия друг другу
- если они не друзья, то это источник скрытого напряжения

# Triadic Closure

- С течением времени таких связей появляется больше



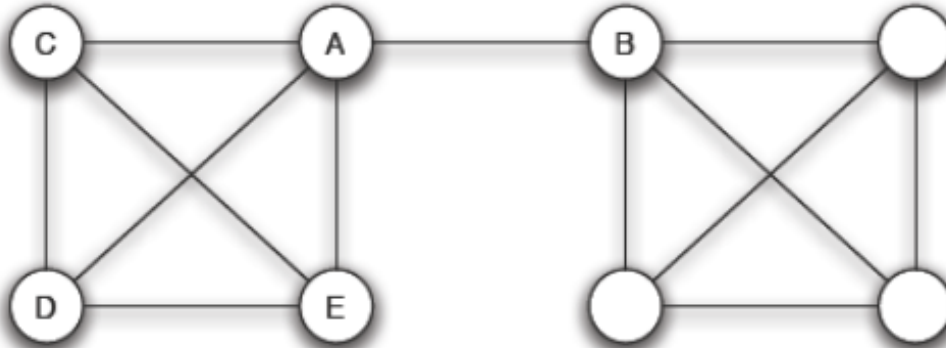
(a) Before new edges form.



(b) After new edges form.

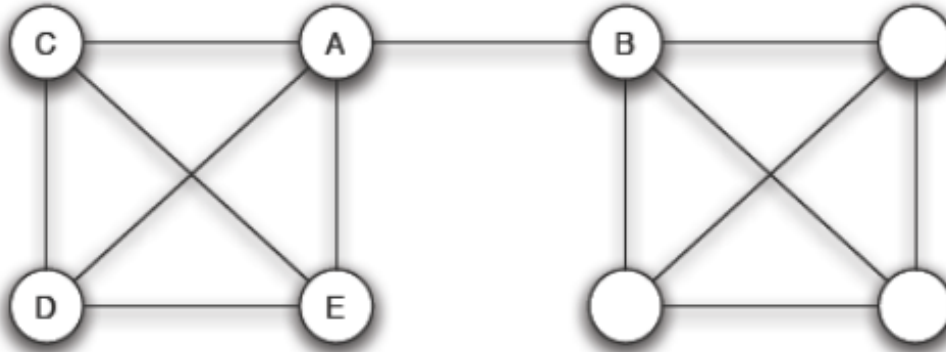
# Мосты

- у А четыре друга
  - С, D, E входят в одну тесно связанную группу
  - В принадлежит к другой части сети -> дает доступ к новым знаниям



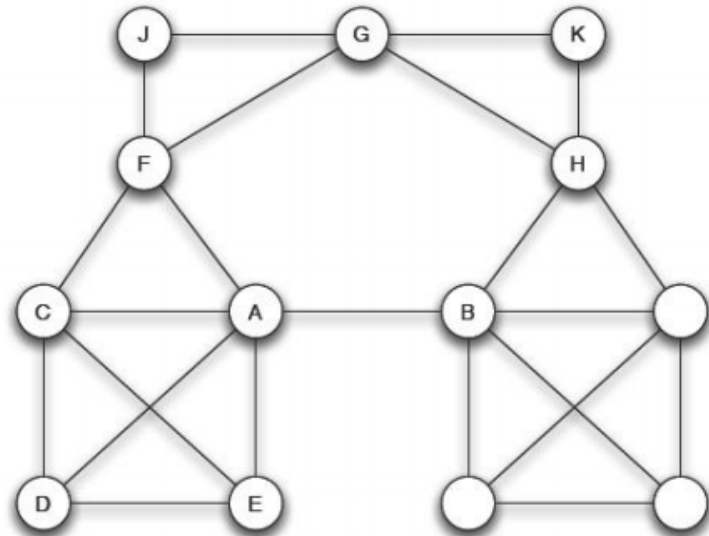
# Мосты

- А-В называется мостом (удаление этой связи приводит к исчезновению связи между группами)



# Локальные мосты

- Локальный мост -- его удаление приводит к увеличению длины пути между частями сети больше, чем на 2 шага



# «Поиск работы»

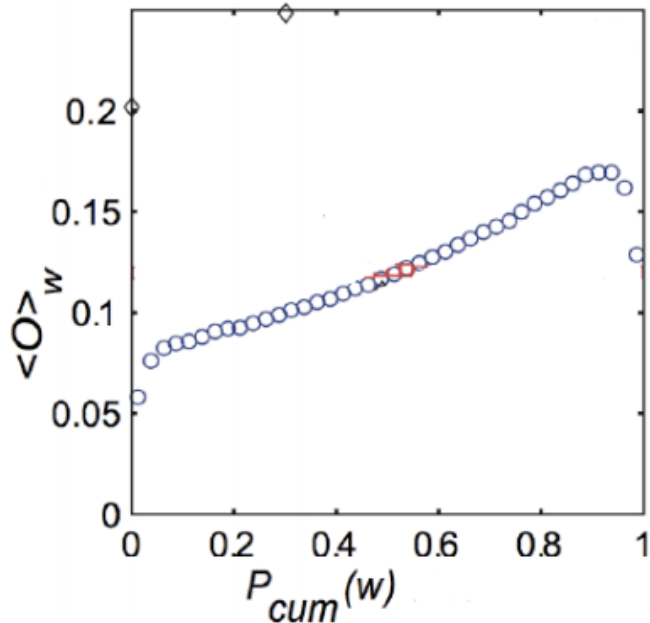
- Локальные мосты соединяют с «полезными» людьми
- Люди из тесно связанной группы имеют примерно одинаковый доступ с одной и той же информации

# Слабые связи и локальные мосты в реальных данных

- Onnela et al.: цифровая коммуникация («Кто с кем разговаривает»)
  - данные мобильных разговоров
  - 20% населения страны
  - наблюдение в течение 18 недель
  - 84% пользователей объединены в связную сеть
- Как измерить силу связи
  - использование времени разговора
  - доля общих соседей (для локальных мостов=0)

# Слабые связи и локальные мосты в реальных данных

- У вершин с меньшим числом общих соседей меньше сила связи




overlap as a function of strength (percentile)



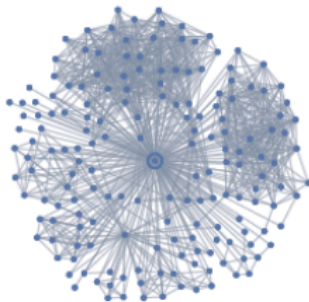
# Сила связей в Facebook

- Cameron and Marlow:

- 
- stronger**
- 3 categories of links (usage over a 1-month period)
- **mutual communication**: user both sent and receive messages from the friend
  - **one-way communication**: user sent messages to the friend (regardless if replied)
  - **maintained relationship**: user followed information of the friend (regardless of messages)
    - “following information”: clicking on content via Facebook’s News Feed service or visiting the friend’s profile
- weaker**

# Сила связей в Facebook

**All Friends**



**Maintained Relationships**



**One-way Communication**

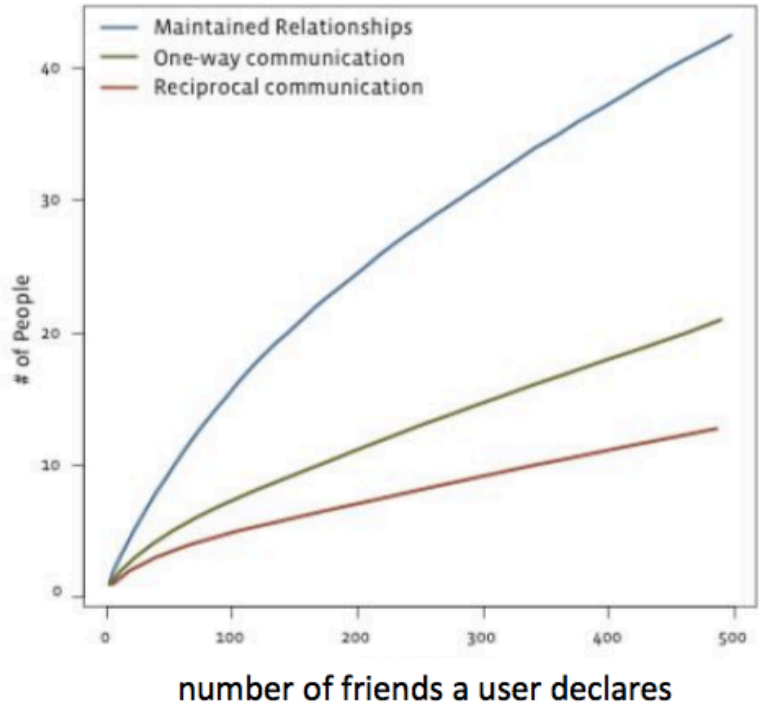


**Mutual Communication**



# Дружба в Facebook

- Декларируемых друзей около 500
- Сильные связи 10-20
- Слабые связи <50



# Число Данбара

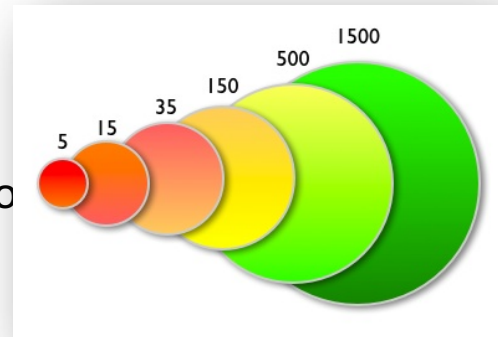
Робин Данбар нашёл связь между размером неокортекса приматов и способностью поддерживать социальные связи

Люди могут комфортно общаться только с ~150 людьми

Это число не включает нефункциональные отношения

При этом эта сеть образуется за счет вложенных слоёв

Поддержание связей требует затрат, пропорциональных слою



# «Глобальная деревня»

Физическое расстояние не играет роли в интернете, соответственно социальное взаимодействие меняется открытостью Всемирной Паутины, простотой с которой люди могут искать онлайн-сообщества по вкусу и взаимодействовать с теми, кто разделяет их интересы и беспокойства

Таким образом, благодаря интернету наш мир --  
“глобальная деревня”

**Так ли это?**

# Интересные ссылки

Dunbar's number and Global Village:

<https://www.youtube.com/watch?v=pjsYPocCqyQ>

<http://www.newyorker.com/science/maria-konnikova/social-media-affect-math-dunbar-number-friendships>

Networks, Crowds, Markets

[www.cs.cornell.edu/home/kleinber/networks-book/](http://www.cs.cornell.edu/home/kleinber/networks-book/)

Учебник по социально-сетевому анализу:

[barabasi.com/book/network-science](http://barabasi.com/book/network-science)

Facebook study:

<https://www.youtube.com/watch?v=OaL7YE4LHew>

Filter Bubble:

[https://www.ted.com/talks/eli\\_pariser\\_beware\\_online\\_filter\\_bubbles?language=en](https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=en)