

Introduction to the Social Web

Recommendation and Mining

Sihem Amer-Yahia

CNRS/LIG

Apr 29th, 2019

Apr 30th, 2019

Yesterday's outline

- Recommender Systems
 - What are recommender systems and how do they work?
 - How are recommender systems evaluated?
- Recommendation challenges
 - Well-known challenges
 - Recommendation diversity
 - Group recommendation
 - Contextual recommendation

Collaborative rating systems

- **Places where users express their opinion on a content item in the form of a rating**
- **Of great interest to:**
 - Analysts who seek to explore users' opinion on items
 - End-users who seek to make choices, find similar/dissimilar users



last.fm



movielens

User data on collaborative rating systems

a set of rating records:

<item attributes, user attributes, rating>

ID	Movie	Name	Gender	Age	Occup.	Rating
r ₁	Toy Story	John	M	young	teacher	4
r ₂	Toy Story	Jennifer	F	old	teacher	3
r ₃	Toy Story	Mary	F	old	teacher	2
r ₄	Titanic	Carine	F	old	other	4
r ₅	Toy Story	Sara	F	young	student	3
r ₆	Toy Story	Martin	M	young	student	5
r ₇	Titanic	Peter	M	young	student	1

Data from MovieLens

MovieLens and IMDb

ID	Title	Genre	Director	Name	Gender	Location	Rating
1	Titanic	Drama	James Cameron	Amy	Female	New York	8.5
2	Schindler's List	Drama	Steven Speilberg	John	Male	New York	7.0

	MovieLens (+IMDb)	BookCrossing
#Users	6,040	38,511
#Items	3,900	260
#Ratings	1,000,209 (million)	196,842
Rating Scale	1 to 5	1 to 10

MovieLens

<http://grouplens.org/datasets/movielens/>

MovieLens 100k

100,000 ratings from 1000 users on 1700 movies.

- [README.txt](#)
- [ml-100k.zip](#)
- [Index of unzipped files](#)

MovieLens 1M

1 million ratings from 6000 users on 4000 movies.

- [README.txt](#)
- [ml-1m.zip](#)

MovieLens 10M

10 million ratings and 100,000 tag applications applied to 10,000 movies by 72,000 users.

- [README.html](#)
- [ml-10m.zip](#)

Social data mining: definition

- **Define social data mining as mining user groups**
 - because labeled user groups exhibit less sparsity and less noise than individual records
 - because labeled groups provide new insights
- **Group: set of rating records describable by a set of attribute values**

Example user groups

- *Young people who rated Woody Allen movies*
- *Middle-aged females in California*
- *People who rated movies starring Scarlett Johansson*
- *Female engineers who rated Star Wars*
- *[25-35] year-old professionals who live in Grenoble and who rated movies starring Sean Penn*

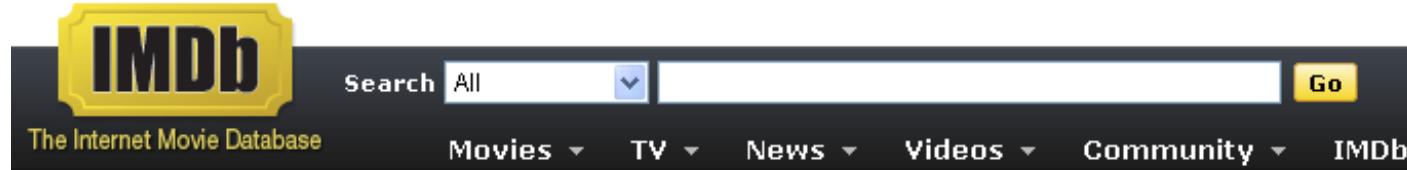
Social data mining problem

Given a rating dataset, discover *good* groups

Challenges and outline

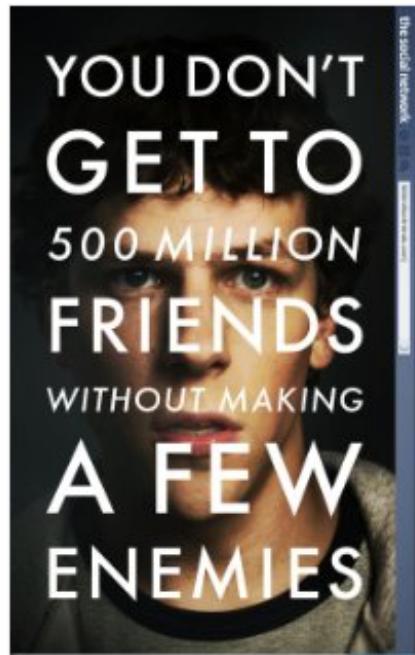
- **Challenges**
 - How to express group quality?
 - How to find groups quickly?
- **Outline**
 1. Social data mining
 2. Multi-objective mining
 3. Open questions
 1. Tag-based mining
 2. Interactive mining

Pre-defined user groups on IMDb



The Internet Movie Database

Movies ▾ TV ▾ News ▾ Videos ▾ Community ▾ IMDb



The Social Network (2010)

PG-13 120 min - [Biography](#) | [Drama](#) - [1 October 2010 \(USA\)](#)



Ratings: 8.0/10 from 146,847 users Metascore: 95/100

Reviews: 522 user

A chronicle of the founding networking Web site.

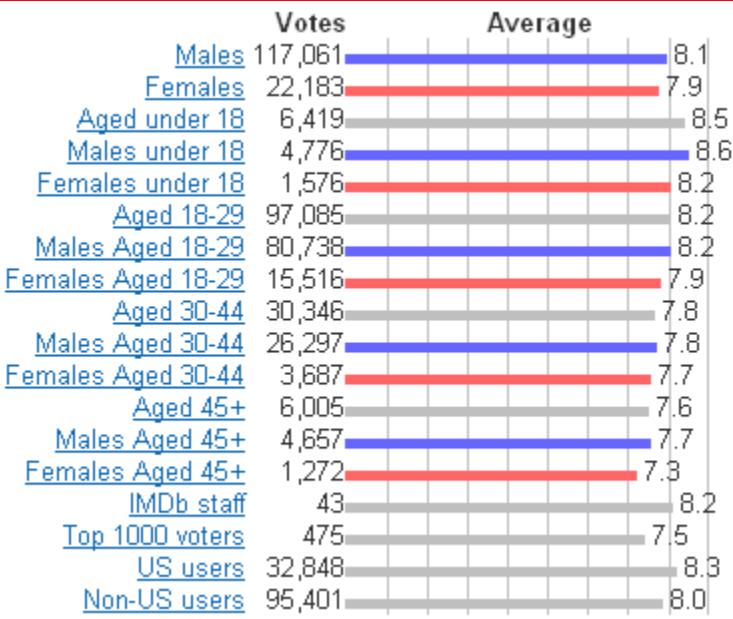
Director: [David Fincher](#)

Writers: [Aaron Sorkin](#) (so)

Stars: [Jesse Eisenberg](#), [A. Timberlake](#)

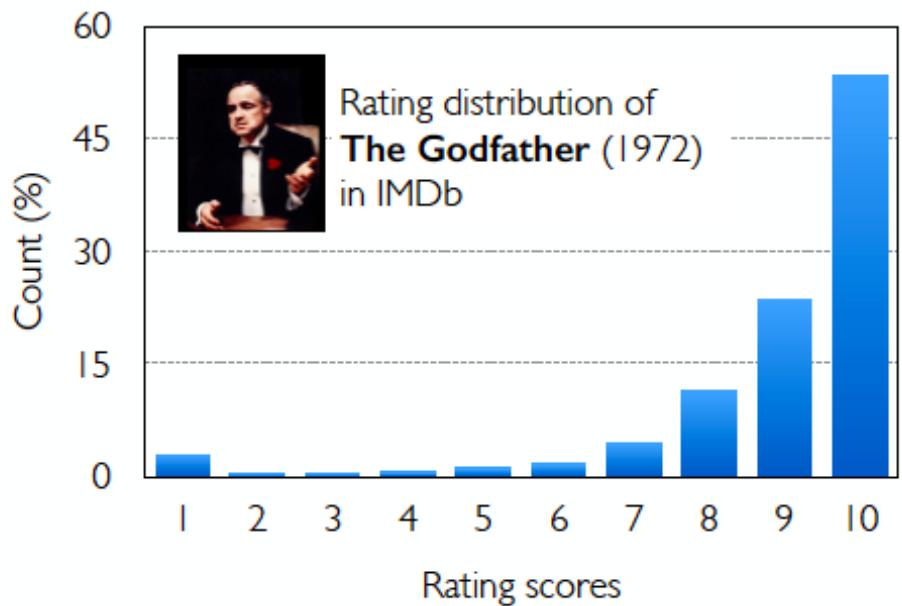
[Watch Trailer](#)

[Add](#)

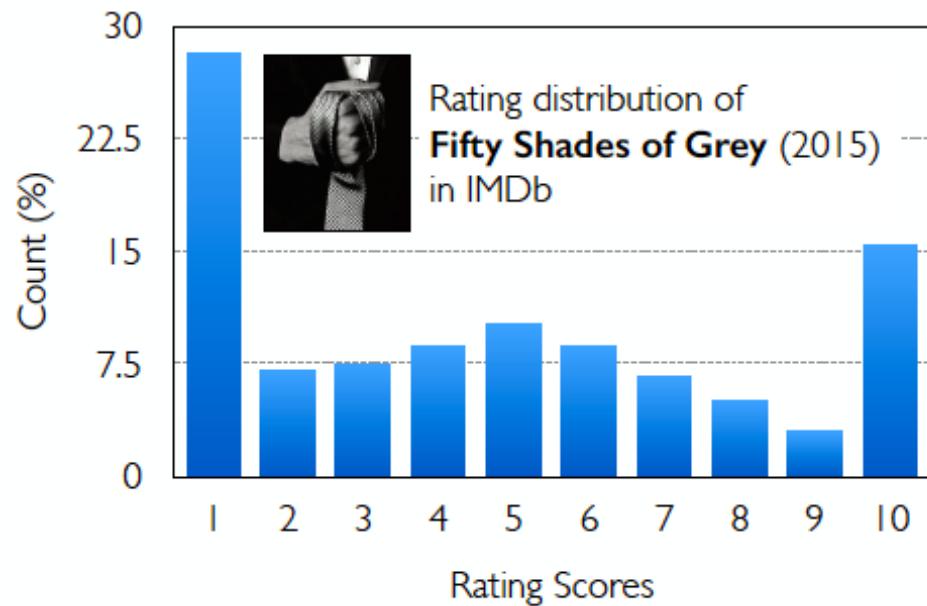


What is a good group, locally speaking?

Homogeneous Rating Distribution



Polarized Rating Distribution



What is a good group, globally speaking?

IMDb



Ratings for
romance
genre movies.



Elena
social
scientist

I believe romantic
movies are mostly
watched by females.

young females
average rating:
3.7

variance:
2

females in DC
average rating:
4.6

variance:
1.5

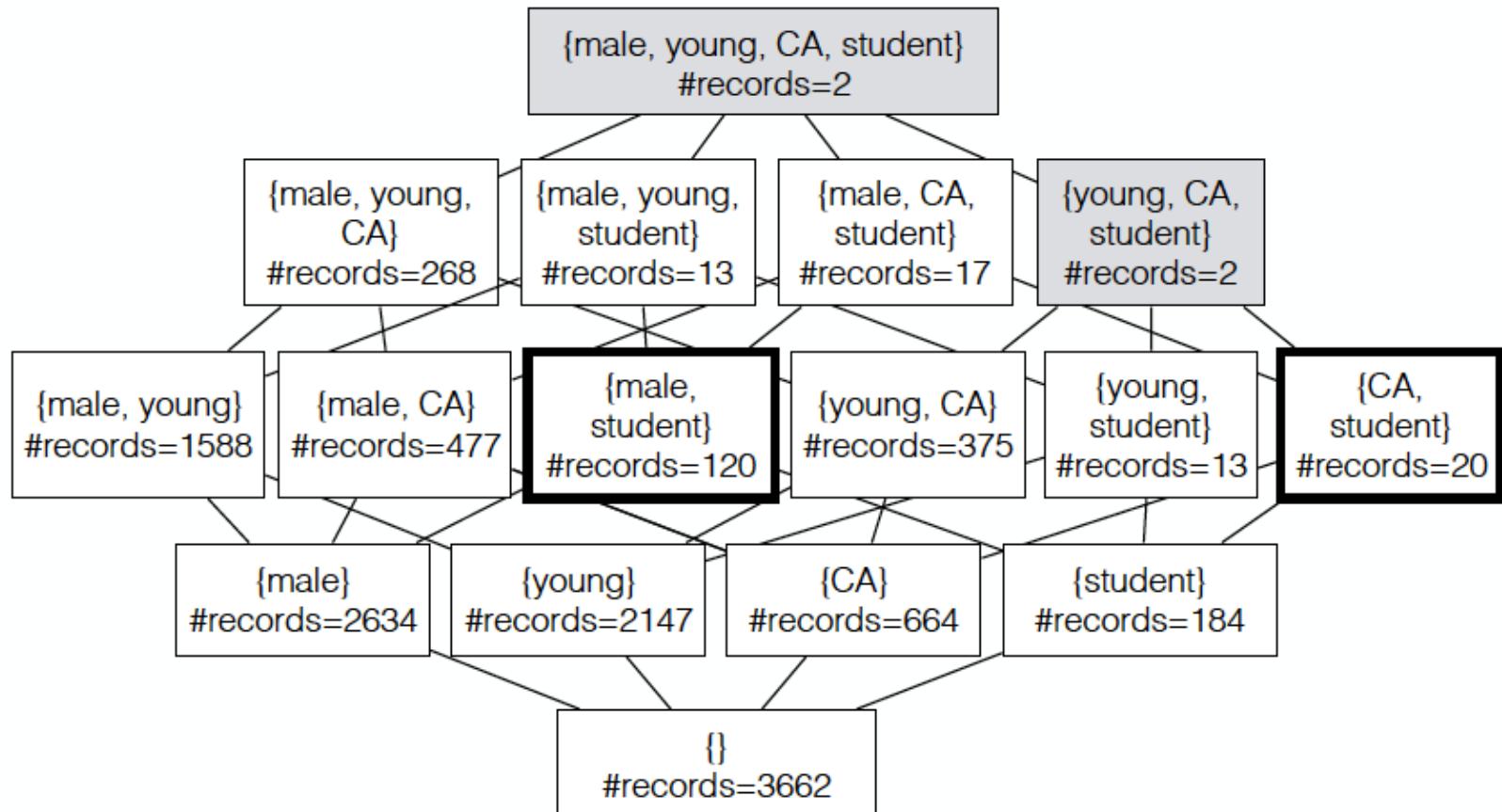
male teenagers
average rating:
3.1

variance:
3.4

— user groups
that cover most
ratings

Partial lattice for the movie Toy Story

An exponential search space



1. An example of user group mining [1]

For an input item covering R_I , ratings, return a set C of k groups, s.t.

description error $\text{error}(C, R_I)$ is minimized, subject to:

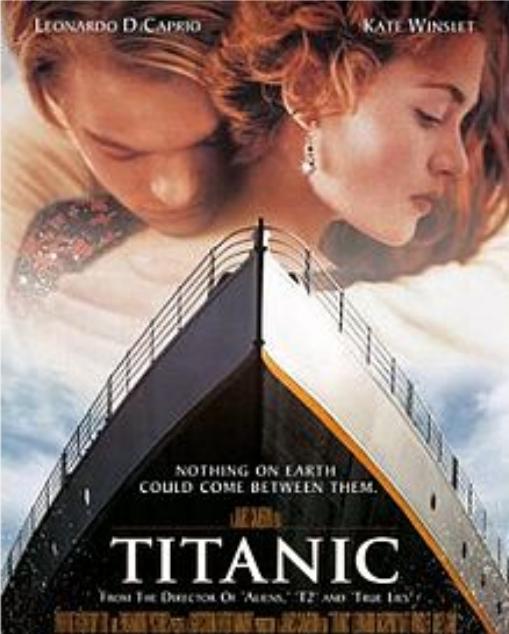
coverage $\text{coverage}(C, R_I) \geq \alpha$

$$\begin{aligned}\text{error}(C, R_I) &= \sum_{r \in R_I} (E_r) \\ &= \sum_{r \in R_I} \text{avg}(|r.s - \text{avg}_{c \in C \wedge r \in c}(c)|)\end{aligned}$$

Mining one group k=1

Identify groups of reviewers who consistently share similar ratings on items

Titanic



LEONARDO DICAPRIO KATE WINSLET

NOTHING ON EARTH
COULD COME BETWEEN THEM.

TITANIC
From the Director Of "Aliens," "T2" And "True Lies"

Titanic (1997)

PG-13 194 min - Adventure | Drama | History - 19 December 1997 (USA)

7.4 Ratings: 7.4/10 from 288,334 users Metascore: 74/100
Reviews: 2,284 user | 174 critic | 34 from Metacritic.com

**Teen-aged female reviewers have rated this movie uniformly
Their average rating: 9.2**

Mining three groups k=3

Identify groups of reviewers who consistently share similar ratings on items

Black Swan



Black Swan (2010)

108 min - [Drama](#) | [Mystery](#) | [Thriller](#) - [17 December 2010 \(USA\)](#)



Ratings: **8.3/10** from 156,148 users Metascore: **79/100**
Reviews: [892 user](#) | [523 critic](#) | [42 from Metacritic.com](#)

Young female reviewers love this movie, average rating: 9.3

Reviewers from New York love this movie, average rating: 8.7

Young male student reviewers hate this movie, average rating: 6.1

User group mining problem

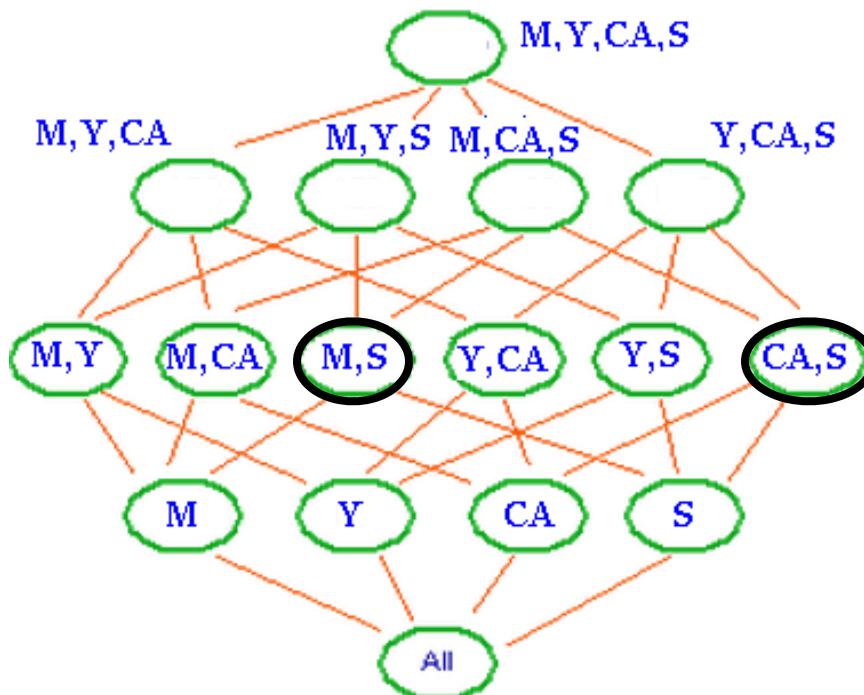
THEOREM 1. *The decision version of the problem of meaningful description mining (DEM) is NP-Complete even for boolean databases, where each attribute ia_j in \mathcal{I}_A and each attribute ua_j in \mathcal{U}_A takes either 0 or 1.*

To verify NP-completeness, we reduce the Exact 3-Set Cover problem (EC3) to the decision version of our problem. EC3 is the problem of finding an exact cover for a finite set U , where each of the subsets available for use contain exactly 3 elements. The EC3 problem is proved to be NP-Complete by a reduction from the Three Dimensional Matching problem in computational complexity theory

Random Restart Hill Climbing Algorithm

$k = 2$

- Satisfy Coverage
- Minimize Error

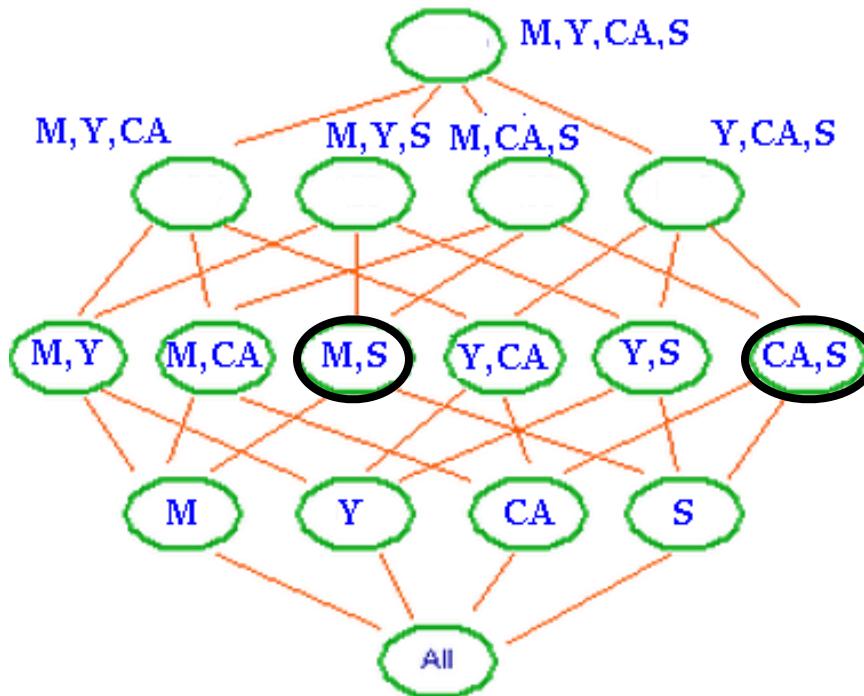


$C = \{\text{Male, Student}\}$
 $\{\text{California, Student}\}$

Random Restart Hill Climbing Algorithm

k = 2

Satisfy Coverage	<input type="checkbox"/>
Minimize Error	<input type="checkbox"/>



C= {Male, Student}
{California, Student}

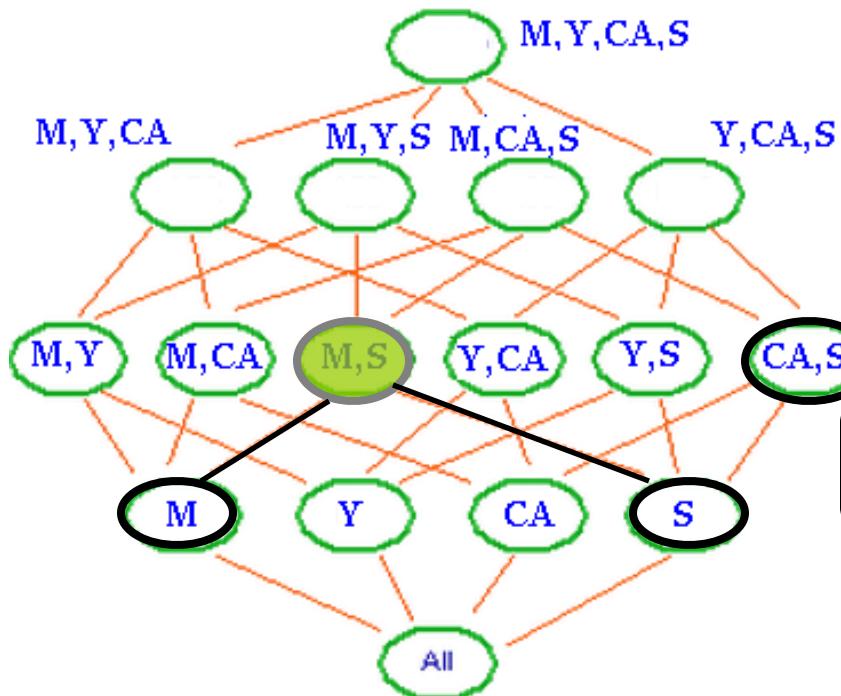
Say, C does not satisfy
Coverage Constraint

Random Restart Hill Climbing Algorithm

k = 2

Satisfy Coverage

Minimize Error



C= {Male, Student}
{California, Student}

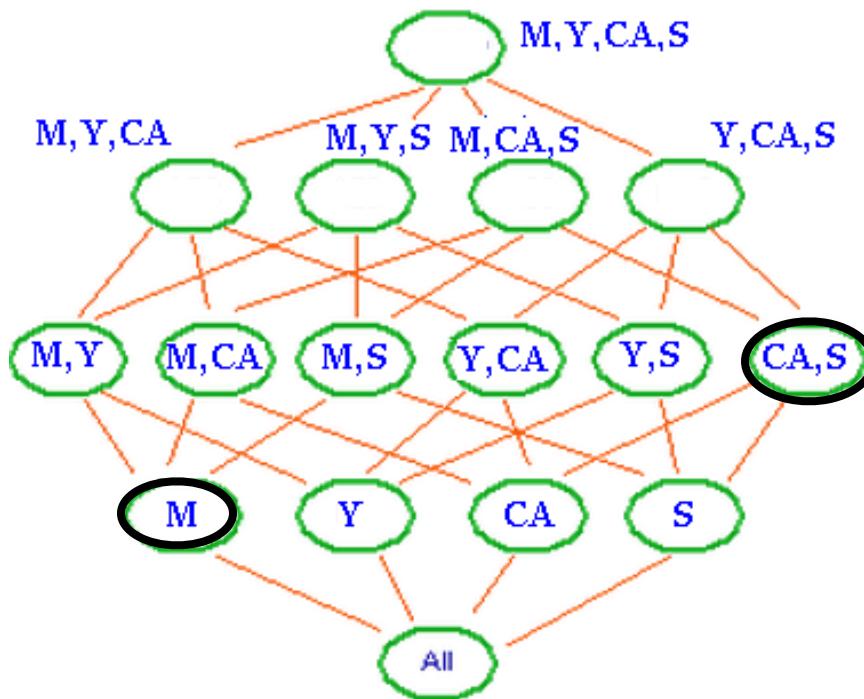
C= {Male}
{California, Student}

C= {Student}
{California, Student}

Random Restart Hill Climbing Algorithm

$k = 2$

Satisfy Coverage	<input checked="" type="checkbox"/>
Minimize Error	<input type="checkbox"/>



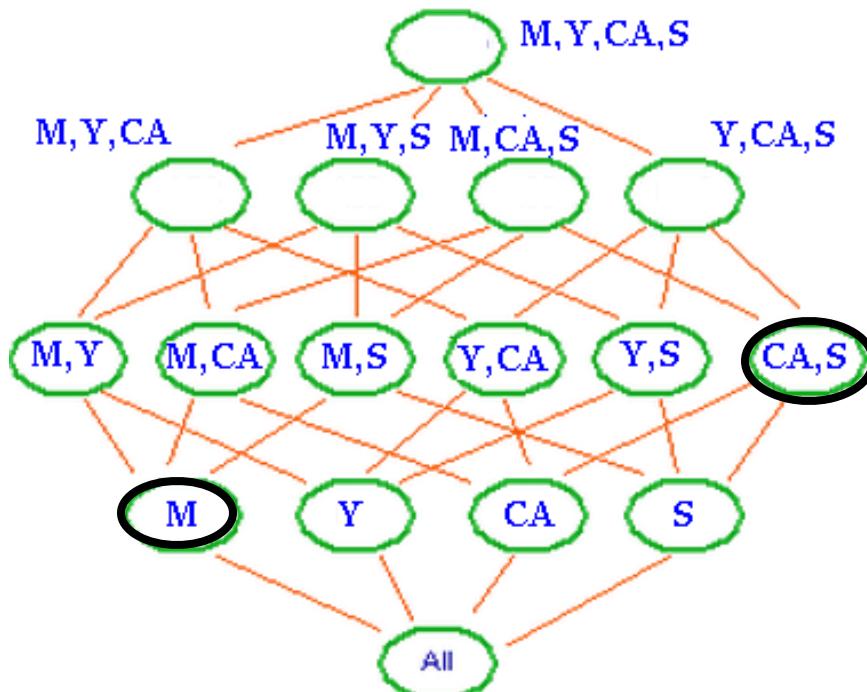
$C = \{\text{Male}\}$
 $\{\text{California, Student}\}$

Say, C satisfies
Coverage Constraint

Random Restart Hill Climbing Algorithm

k = 2

Satisfy Coverage	<input checked="" type="checkbox"/>
Minimize Error	<input type="checkbox"/>



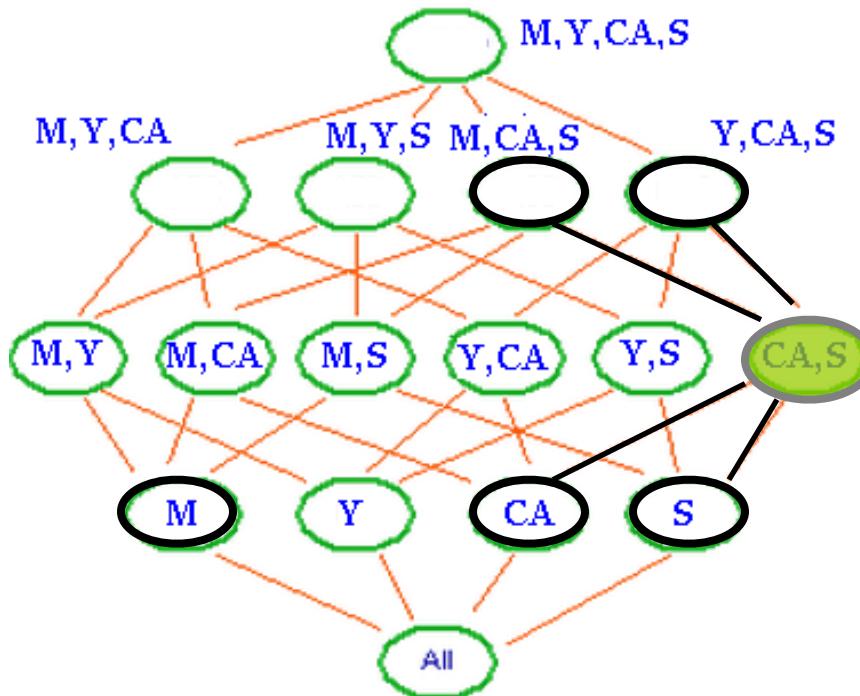
C= {Male}
{California, Student}

Random Restart Hill Climbing Algorithm

k = 2

Satisfy Coverage

Minimize Error

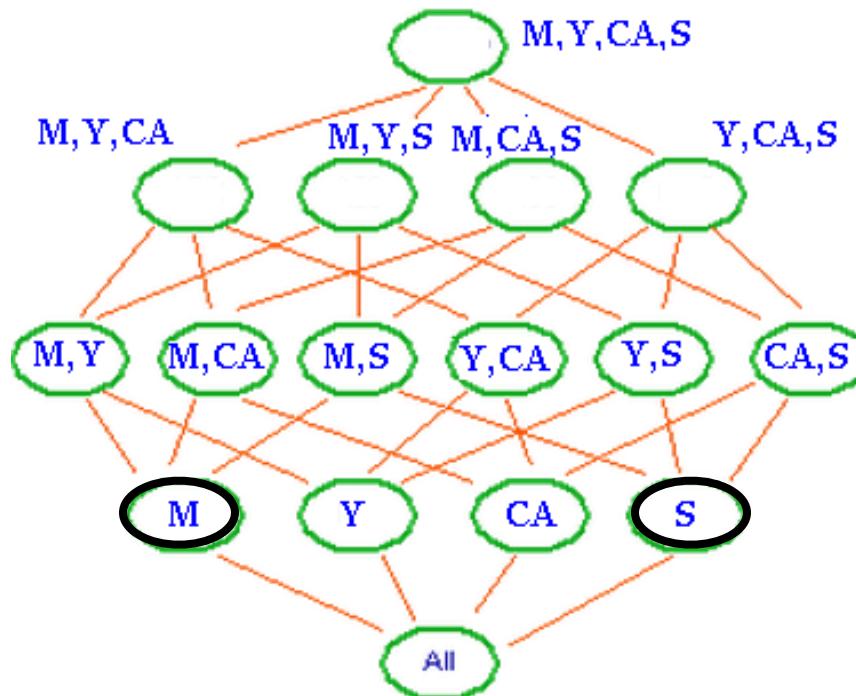


C= {Male}
{California, Student}

Random Restart Hill Climbing Algorithm

$k = 2$

- Satisfy Coverage
- Minimize Error



$C = \{\text{Male}\}$
 $\{\text{Student}\}$

2. Another example of user group mining

Identify groups of reviewers who consistently disagree on item ratings

Schindler's List



Schindler's List (1993)

R 195 min - [Biography](#) | [Drama](#) | [History](#) - [15 December 1993 \(USA\)](#)

 8.9 Ratings: 8.9/10 from 329,773 users Metascore: 93/100
Reviews: 959 user | 95 critic | 23 from Metacritic.com

Teen-aged female reviewers and male middle-aged reviewers have rated this movie inconsistently; their average rating: 7.5

- Middle-aged male reviewers love this movie, their average rating: 9.1
- Teen-aged female reviewers hate this movie, their average rating: 6.2

3. A third example of user group mining People like/unlike me

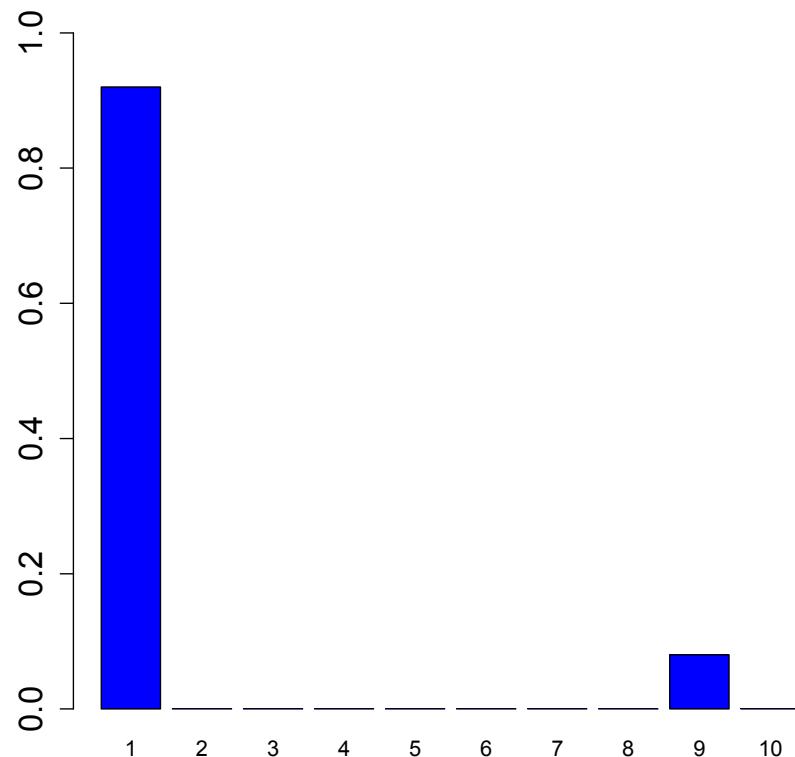
Mary: 32 years, live in Bethlehem, USA
dislikes books by Debbie Macomber



- ✓ 25 middle-aged people, live in the **USA**, dislike "204 Rosewood Lane"
- ✗ 11 people, live in the **USA**, like "Changing Habits"

We call this a rating map.

Mary's distribution for Debbie Macomber's books



EMD as a rating comparison measure

$$\rho_1 = [0.9, 0.025, 0.025, 0.025, 0.025]$$

$$\rho_2 = [0.025, 0.9, 0.025, 0.025, 0.025]$$

$$\rho_3 = [0.025, 0.025, 0.025, 0.025, 0.9]$$

Measure	(ρ_1, ρ_2)	(ρ_1, ρ_3)
Cosine	0.058	0.058
KL-Divergence	3.13	3.13
JS-Divergence	0.53	0.53
Euclidean distance	1.24	1.24
Hellinger Distance	0.791	0.791
Total Variation Distance	0.875	0.875
Renyi Entropy Distance (0.5 order)	1.962	1.962
Battacharya Distance	0.981	0.981
Distance correlation	0.2500	0.2500
Signal Noise Ratio	2.0372	4.221
Lukaszyk-Karmowski Metric	1.1625	3.525
EMD	0.875	3.5

People like me/unlike me problem [2]

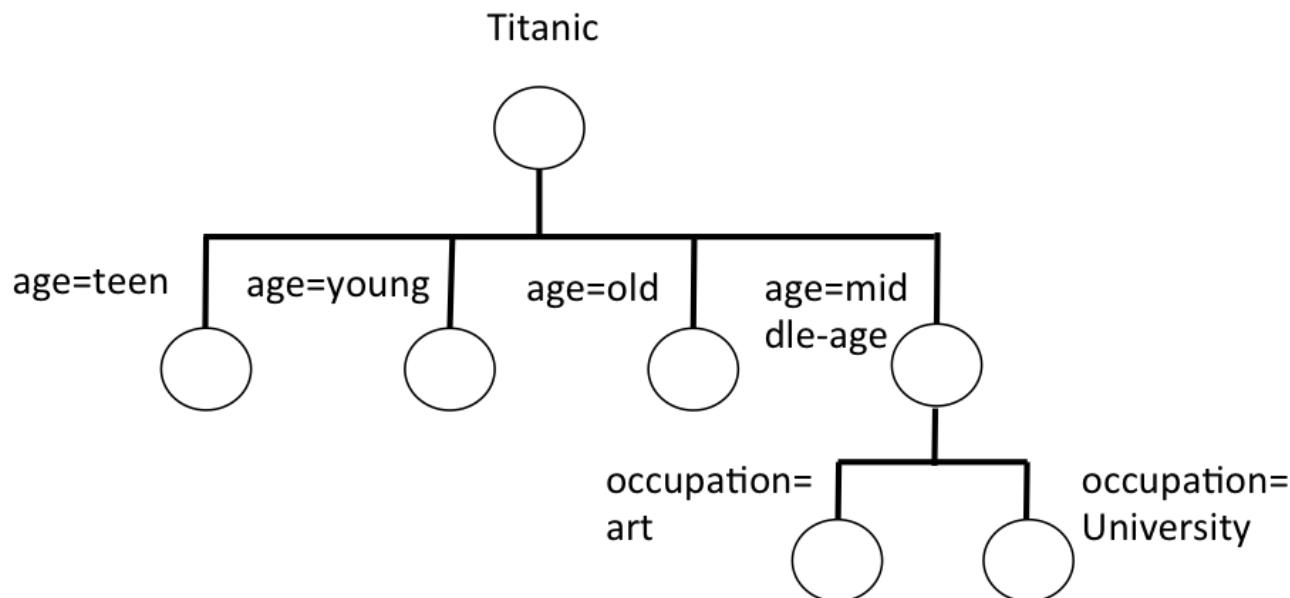
Given a set of input distributions $\{\rho_1, \dots, \rho_k\}$, find the largest *distinct user groups* whose distribution is the closest to one of $\{\rho_1, \dots, \rho_k\}$ (using an *EMD threshold* θ)

- Groups with a shorter description are preferred
- Large groups are preferred
- Groups with different descriptions are preferred

[2] *Exploring Rated Datasets with Rating Maps* Sihem Amer-Yahia, Sofia Kleisarchaki, Naresh Kumar Kolloju, Laks V.S. Laskhmanan, Ruben H. Zamar
WWW 2017

Partition Decision Tree (PDT)

- The set of rating records can be organized in a PDT
- Each node is a user group with a description



Brief sketch of algorithms

- **DTAlg**

- minimizes description length by finding a minimum height partition decision tree
- classic decision trees driven by gain functions like entropy and gini-index.

$$\text{Gain}(\text{Attr}_i) = \frac{n}{\sum_{j=1}^n \min_{\rho \in \{\rho_1, \dots, \rho_k\}} \text{EMD}(c_j^i, \rho)}$$

- **Random Forests**

- splitting input dataset hurts coverage
- runs multiple iterations of DTAlg with different splitting attributes and combines trees with RF-Cluster, RF-Desc, RF-Random, RF-Size, and RF-EMD

Summary and outline

1. Social data mining

- formulated as mining user groups
- a constrained optimization problem with a single objective:
 - groups whose ratings are uniform/polarized, and that satisfy a coverage constraint
 - largest groups whose ratings are close to an input distribution
- hard problems that necessitate appropriate heuristics

2. Multi-objective mining

3. Open questions

Summary and outline

1. Social data mining

- formulated as mining user groups
- a constrained optimization problem with a single objective:
 - groups whose ratings are uniform/polarized, and that satisfy a coverage constraint
 - largest groups whose ratings are close to an input distribution
 - hard problems that necessitate appropriate heuristics

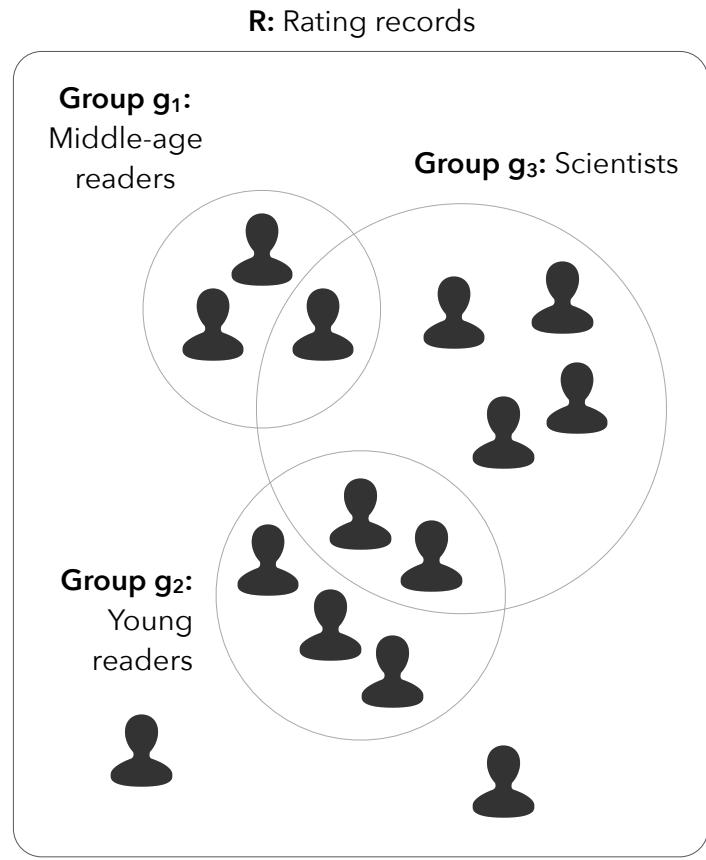
2. Multi-objective mining [3]

3. Open questions

User Group Quality Dimensions

For a group-set G and set of rating records R

- › Coverage. The percentage of ratings in R contained in groups in G.
 - › $\text{coverage}(G, R) = | \cup_{g \in G} (r \in G, u \in R) | / |R|$
- › Diversity. Measures how distinct groups in group-set G are from each other.
 - › $\text{diversity}(G, R) = 1 / (1 + \sum_{g, g' \in G} |r \in R, r \in g, r \in g'|)$
- › Variance. Average diameter of rating scores.
 - › $\text{diameter}(G) = \text{avg}_{g \in G} (\max_{r \in g} (r.s) - \min_{r' \in g} (r'.s))$



Problem Definition

- A multi-objective optimization problem
- Given set of ratings R , identify all group-sets where each group-set G satisfies:
- Objectives
 - $\text{coverage}(G, R)$ is maximized;
 - $\text{diversity}(G, R)$ is maximized;
 - $\text{diameter}(G, R)$ is minimized;
- Constraints
 - $|G| \leq k$;
 - $\forall g \in G: |g| \geq \sigma$.
- Proved to be NP-Complete by a reduction from the Exact 3-Set Cover problem (EC3)

Multi-Objective Optimization Principles

Definition 2 (Plan). *Plan p_i , associated to a group-set G_i for a set of rating records $R \subseteq \mathcal{R}$, is a tuple $\langle |G_i|, \text{coverage}(G_i, R), \text{diversity}(G_i, R), \text{diameter}(G_i) \rangle$.*

Definition 3 (Sub-plan). *Plan p_i is the sub-plan of another plan p_j if their associated group-sets satisfy $G_i \subseteq G_j$.*

Definition 4 (Dominance). *Plan p_1 dominates p_2 if p_1 has better or equivalent values than p_2 in every objective. The term “better” is equivalent to “greater” for maximization objectives (e.g., diversity, coverage and polarization), and “lower” for minimization ones (e.g., homogeneity). Furthermore, plan p_1 strictly dominates p_2 if p_1 dominates p_2 and the values of objectives for p_1 and p_2 are not equal.*

Definition 5 (Pareto Plan). *Plan p is Pareto if no other plan strictly dominates p . The set of all Pareto plans is denoted as \mathcal{P} .*

NP-Completeness

Theorem 1. *The decision version of our problem is NP-Complete.*

Proof. (sketch) It is shown in [4] that a single-objective optimization problem for user group discovery is NP-Complete by a reduction from the Exact 3-Set Cover problem (EC3). There, homogeneity is maximized and a threshold on coverage is satisfied. In our case, two new conflicting dimensions (diversity and coverage) are added. This means that the problem in [4] is a *special case* of ours, hence our problem is obviously harder. \square

Optimality Principles

- Optimality Principle (POO). In case of maximization, if the objective value(s) of sub-plans of a plan p increases, then the objective value(s) of p cannot decrease.
- Near-optimality Principle (PONO). Given an objective f and $\alpha \geq 1$, derive G' from G by replacing G_1 by G'_1 and G_2 by G'_2 . Then $f(G'_1) \geq f(G_1) \times \alpha$ and $f(G'_2) \geq f(G_2) \times \alpha$ together imply $f(G') \geq f(G) \times \alpha$.
 - Proofs are available in our technical report

Near-Optimality Principle

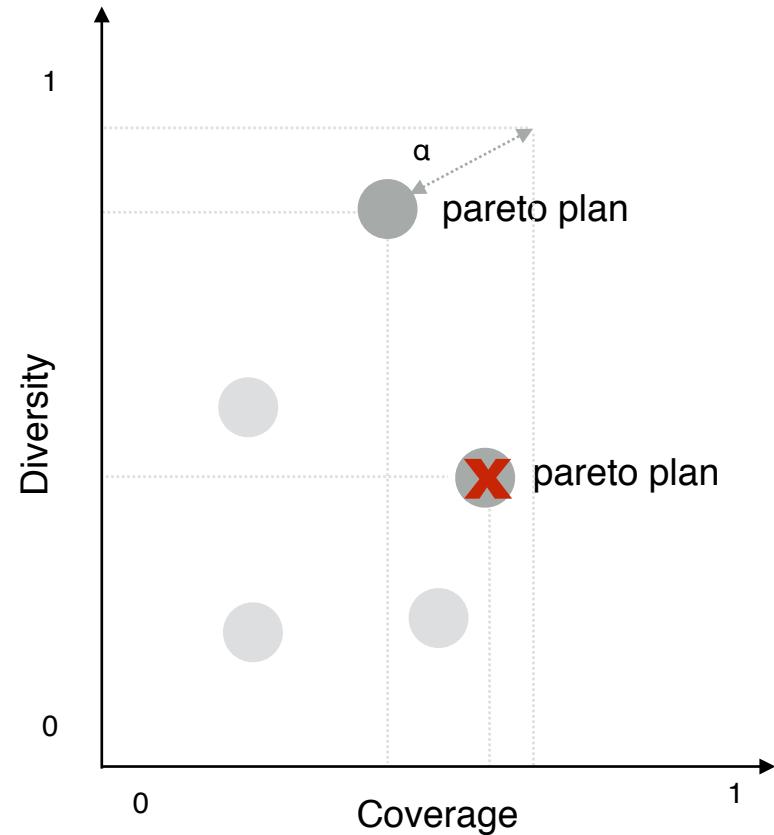
Definition 6 (PONO). *Given a maximization objective f (e.g., diversity, coverage, polarization) and $\alpha \geq 1$, let p_1 be a plan with sub-plans p_{11} and p_{12} . Derive p_2 from p_1 by replacing p_{11} by p_{21} and p_{12} by p_{22} where p_{21} and p_{22} are sub-plans of p_2 . Then $f(G_{21}) \geq f(G_{11}) \times \alpha$ and $f(G_{22}) \geq f(G_{12}) \times \alpha$ together imply $f(G_2) \geq f(G_1) \times \alpha$. The extension for a minimization objective is straightforward.*

Definition 7 (Approximated Dominance). *Let $\alpha \geq 1$ be the precision value, a plan p_1 α -dominates p_2 if for every maximization objective f (e.g., diversity, coverage, polarization), $f(G_1) \geq f(G_2) \times \alpha$. The extension for a minimization objective is straightforward.*

Definition 8 (Approximated Pareto Plan). *For a precision value α , plan p is an α -approximated Pareto plan if no other plan α -dominates p .*

Multi-Objective Group Discovery

- Set of optimal solutions as Pareto plans.
- Exhaustive execution to verify all group-sets of size 1 to k.
- Generating fewer plans makes a Multi-Objective optimization algorithm run faster.



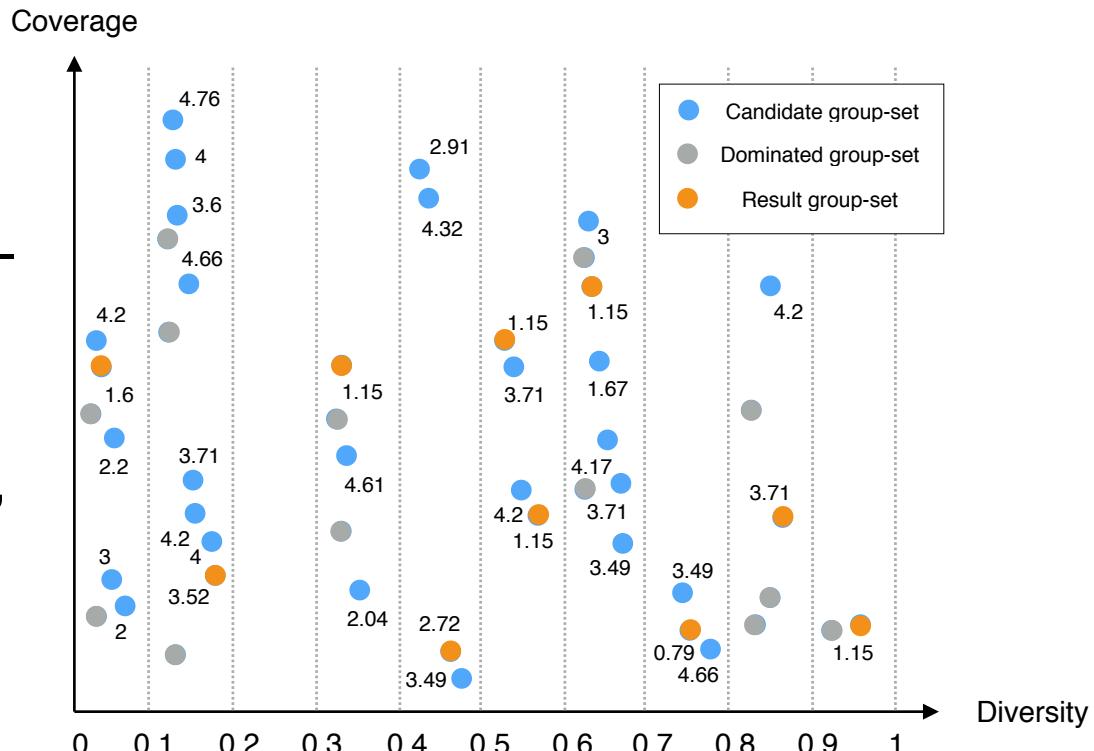
Approximation Algorithm

- Input. $k, \alpha > 1, R$
 - Output. Pareto result set P
1. $P \leftarrow \emptyset$
 2. For all user groups g do
 - 3. $G \leftarrow$ Singleton group-set containing g
 - 4. If G is not α -dominated by any other group-set $\in P$, then add G to P
 5. end
 6. For $n \in [2, k]$ do
 - 7. For each possible group-set G of size n do
 - 8. If G is not α -dominated by any other group-set $\in P$, then add G to P
 - 9. End
 10. End
 11. Return P

— Based on: Trummer, I., Koch, C.: Approximation schemes for many-objective query optimization. In SIGMOD 2014

Heuristic Algorithm

1. Find n group-sets with maximized coverage
 2. Remove dominated group-sets in each diversity interval.
 3. For each diversity interval, report the group-set with the optimal value of diameter.



Approximation vs Heuristic

Solutions

Approximation beats heuristic in at least 62% of solutions.



Approximation algorithm generates much more solutions than heuristic.

Objectives

Heuristic achieves around 50% of supremacy.



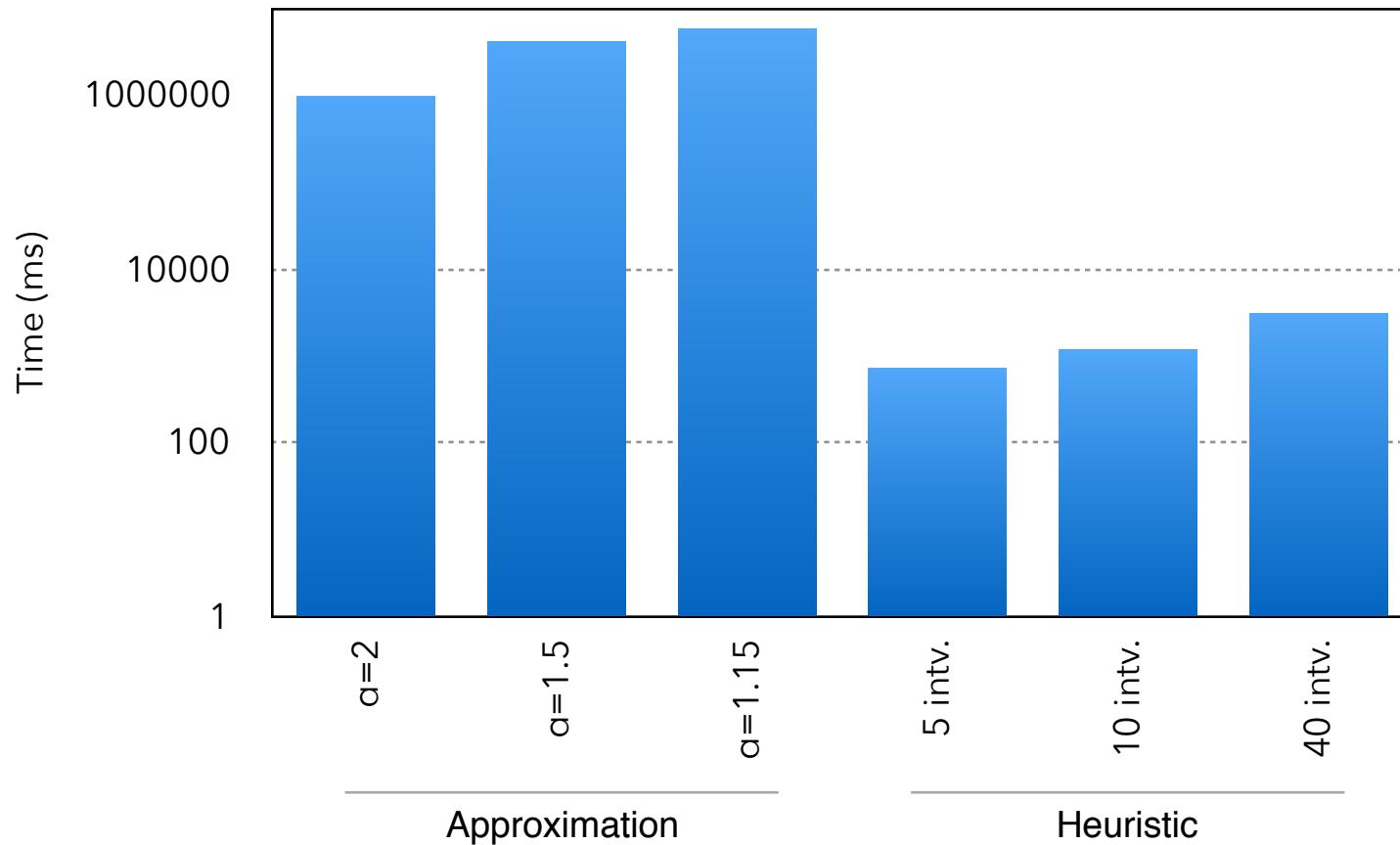
Heuristic can return an acceptable representative subset of results.

Choosing between Algorithms

- Approximation algorithm in an offline context to produce an exhaustive set of user groups with a precision defined by α for further analysis.
- Heuristic algorithm in an online or streaming context, to immediately produce a representative subset of results.

Experiments: Performance

- MovieLens dataset for the movie American Beauty



Outline

1. Social data mining
2. Multi-objective mining
3. Open questions
 1. **Tag-based mining**
 2. Interactive mining

Collaborative tagging system (Amazon)

amazon.com

Hello. [Sign in](#) to get personalized recommendations. New customer? [Start here.](#)

Your Amazon.com | Today's Deals | [Gifts & Wish Lists](#) | [Gift Cards](#)

Shop All Departments Search Electronics

Camera & Photo All Electronics Brands Bestsellers Digital SLRs & Lenses Point-and-Shoots Camcorders

Nikon Coolpix L22 12.0MP Digital Camera with 3.6x Optical Zoom and 3.0-Inch LCD (Red-primary)
by [Nikon](#)

 ★★★★★ (450 customer reviews) |  (94)

Price: **\$79.99**

Tags Customers Associate with This Product ([What's this?](#))
Click on a tag to find related items, discussions, and people.

Check the boxes next to the tags you consider relevant or enter your own tags in the field below.

- [nikon coolpix l22](#) (64) [gift](#) (3)
- [nikon coolpix](#) (47) [lightweight](#) (2)
- [digital camera](#) (33) [12mp](#) (1)
- [nikon](#) (32) [average](#) (1)
- [point_and_shoot](#) (23) [avi video](#) (1)
- [cheap](#) (11) [bad nikon](#) (1)
- [five_star](#) (11) [cool price for an excellent product](#) (1)
- [aa batteries](#) (10) [crappy camera](#) (1)
- [easy_carry_camera](#) (4) [great value](#) (1)
- [affordable](#) (3) [zoom](#) (1)
- [lcd](#) (1)
- [many_photo_settings](#) (1)
- [poor_customer_service](#) (1)
- [camcorder](#) (1)
- [teen](#) (1)
- [underwater_digital_camera](#) (1)
- [unreliable](#) (1)
- [user-friendly](#) (1)

Collaborative tagging system (LastFM)

last.fm

Music Radio Events Charts Community

Join Login

Help Last.fm's scientists with music research »

Artist Biography Pictures Videos Albums Tracks Events News

Music » Adele » Tracks » Rolling In The Deep

Adele – Rolling In The Deep (3:46)
On 5 albums [see all](#)
Buy at Amazon MP3 (\$1.29) | Send Ringtones to Cell
[More options](#) [Save](#)

Popular tags: soul, pop, female vocalists, adele, british [See more](#)

Shouts: 767 shouts

Share this track:

[Send](#) [Tweet](#) [Recommend](#) 259

Tags

00s 10s 2010s **adele** adult alternative alternative amazing voice asdf awesome beautiful beautiful track best of 2011 bittersweet blues breakup brilliant lyrics brilliant **british** chill cool do you want the truth or something beautiful favorites favourite female vocalist **female vocalists** female vocals fossa fucking awesome fucking genius german number 1 gokyel tune hand claps heartbreak i can play this on guitar i wish i wrote this song indie rock instant goosebumps jazz legendary love **love at first listen** neo-soul nice instrument perfect piano piano rock **pop** pop rock power song powerful pure magic relaxing rolling in the deep singer-songwriter **soul** soulful soundtrack of my life stuck in my head taught me to grow 2011

Track Stats

3,477,957 Scrobbles | 314,464 Listeners

Recent Listening Trend



MovieLens instances (with tags) [4]

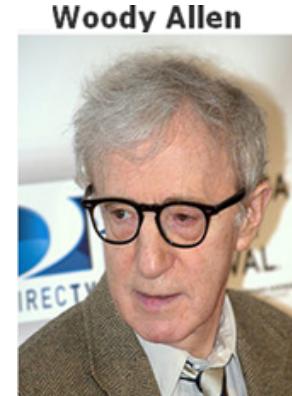
ID		Title	Genre	Director		Name	Gender	Location		Tags
1		Titanic	Drama	James Cameron		Amy	Female	New York		love, Oscar
2		Schindler's List	Drama	Steven Speilberg		John	Male	New York		history, Oscar

[4] Mahashweta Das, Saravanan Thirumuruganathan, Sihem Amer-Yahia, Gautam Das, Cong Yu. Who Tags What? An Analysis Framework. Proceedings of the VLDB Endowment (PVLDB), VLDB Endowment, 2012, 5 (11), pp.1567-1578.

Exploring collaborative tagging in MovieLens

fetish Greek
composers London time
disturbing luck immoral
terrible Dostoyevsky Napolean Miramax Gersin
matchpoint leftovers play comedy tragedy
dramatic Howard-Cosell
Woody Allen New-York
magick wheat Noiva-Nervosa
favorite brilliant Hollywood student person Russian
version ulcer classic cinema ring Lolita

Tag Signature for all Users

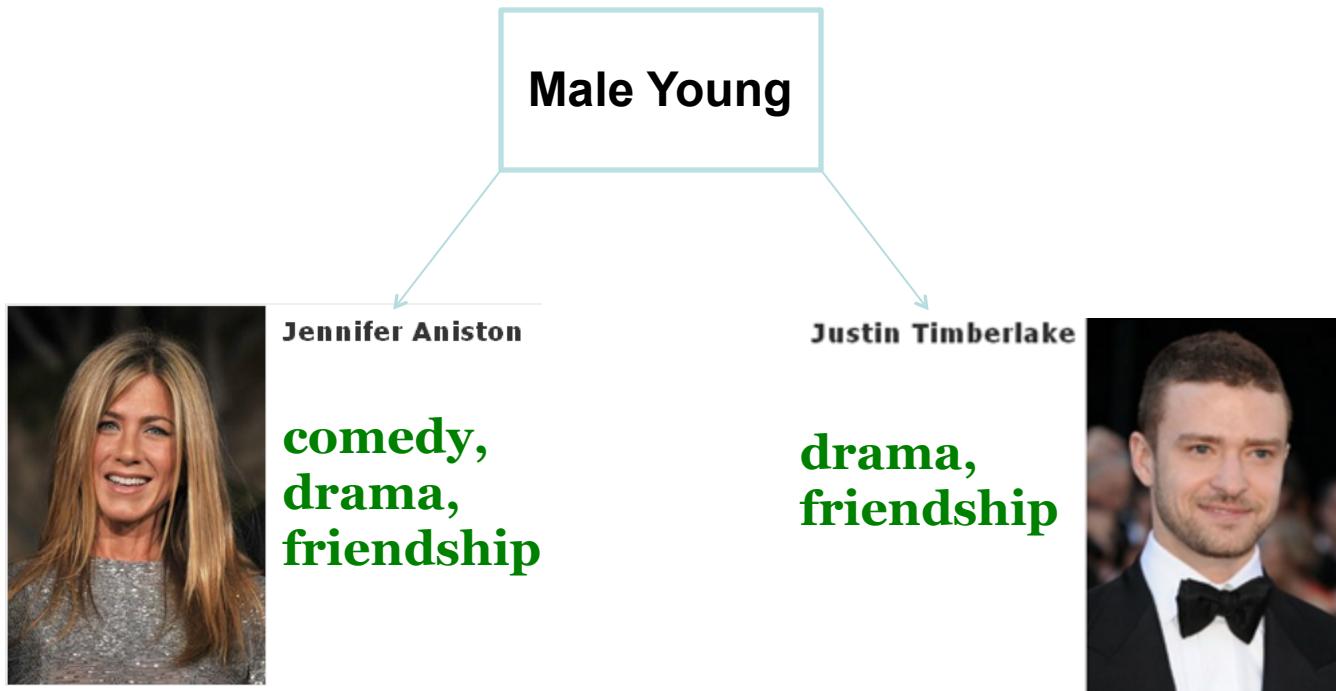


luck
investigations
Miramax wealth artists
Broadway
student
favorite
magick
brilliant
madness
tennis
Barcelona
threesome person
fake playwright
documentary
1930s
Dostoyevsky
terrible
Woody Allen play
murder affair
futuristic
classic psychiatry
mobster

Tag Signature for all CA Users

Exploring collaborative tagging in MovieLens

Identify similar groups of reviewers who share similar tagging behavior for different items



Exploring collaborative tagging in MovieLens

Identify diverse groups of reviewers who have different tagging behavior for the same items

Male Teen
gun, special effects

Female Teen
violence, gory

Genre: Action

Most Popular Action Feature Films

1.  [Prometheus](#) (2012)
 7.9/10
A team of explorers discover a clue to the origins of mankind on Earth, leading them on a journey to the darkest corners of the universe. There, they must fight a terrifying battle to save the future of the human race.
Dir: Ridley Scott With: Noomi Rapace, Logan Marshall-Green, Michael Fassbender
Action | Horror | Sci-Fi 124 mins. 
2.  [The Avengers](#) (2012)
 8.6/10
Nick Fury of S.H.I.E.L.D. brings together a team of super humans to form The Avengers to help save the Earth from Loki and his army.
Dir: Joss Whedon With: Robert Downey Jr., Chris Evans, Scarlett Johansson
Action | Adventure | Sci-Fi 143 mins. 
3.  [Snow White and the Huntsman](#) (2012)
 6.7/10
In a twist to the fairy tale, the Huntsman ordered to take Snow White into the woods to be killed winds up becoming her protector and mentor in a quest to vanquish the Evil Queen.
Dir: Rupert Sanders With: Kristen Stewart, Chris Hemsworth, Charlize Theron
Action | Adventure | Drama | Fantasy 127 mins. 

Outline

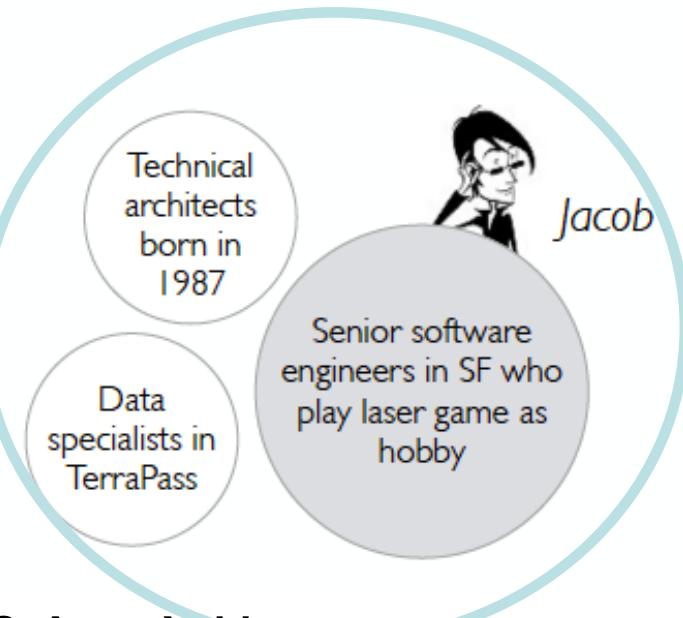
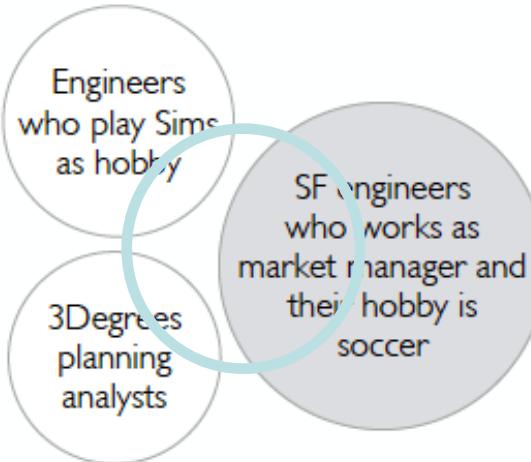
1. Social data mining
2. Multi-objective mining
3. Open questions
 1. Tag-based mining
 - 2. Interactive mining**

Interactive mining [5]

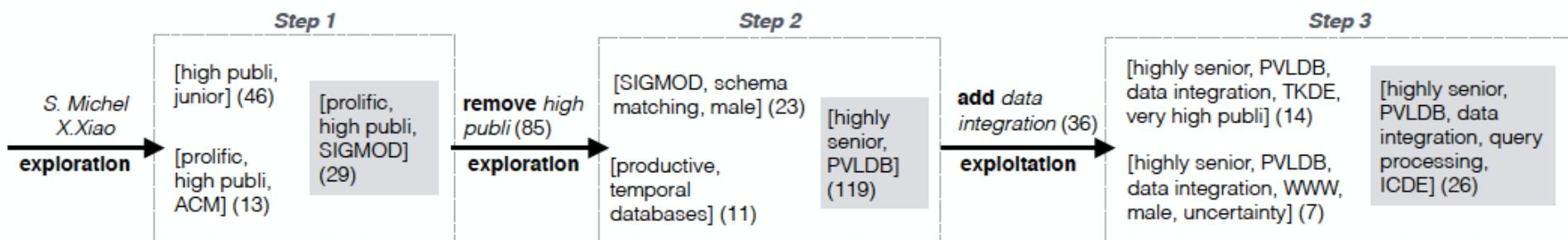


I met a guy in last night party in San Francisco (SF) but lost his phone number and I don't remember his name! I only remember that he works as engineer.

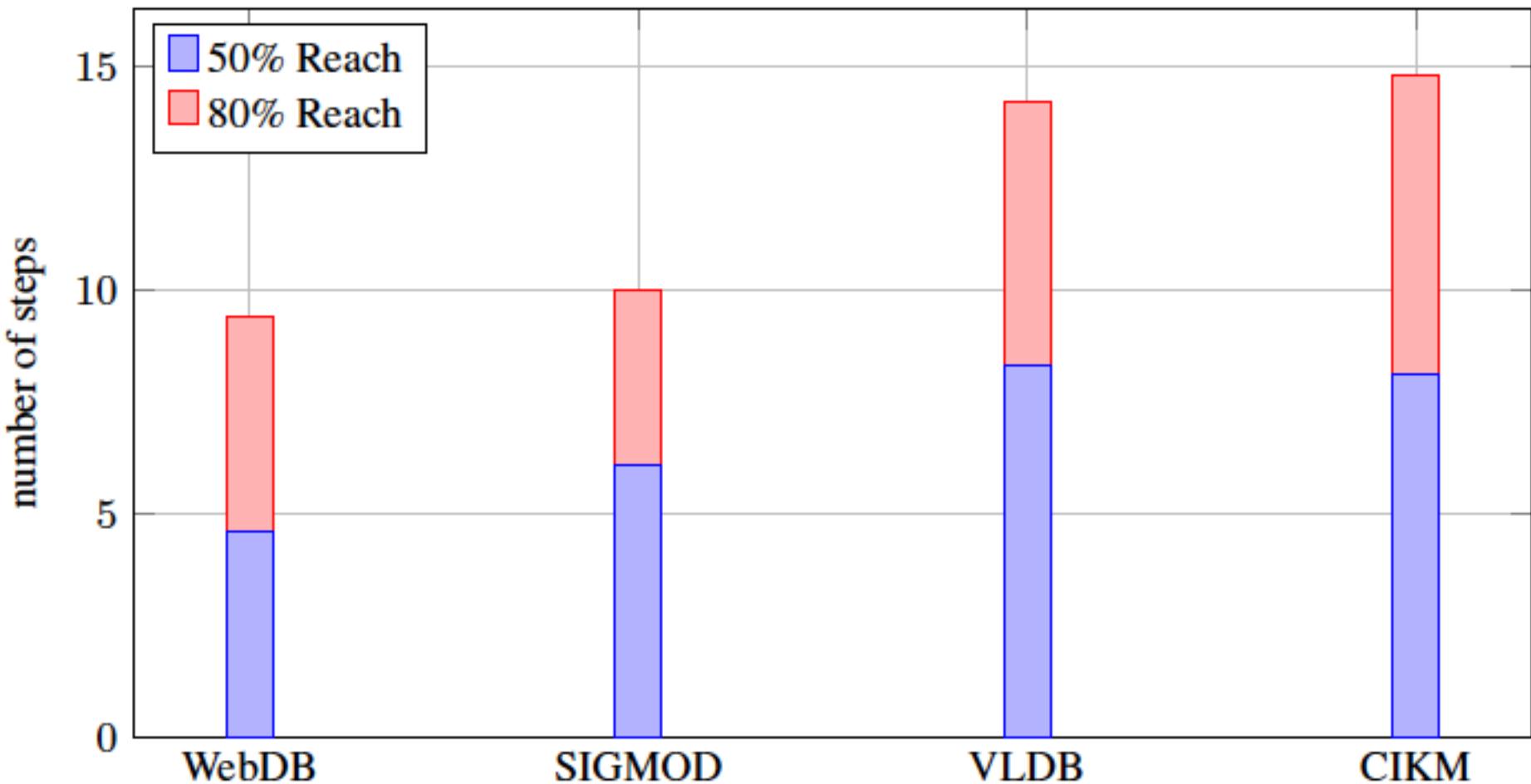
Engineers in SF



Interactive mining



Interactive mining



Open questions with interactive mining

- **Personalized mining [6]**
 - learn a model and suggest mining operations
- **A benchmark for evaluating interactive exploration**
 - Generic: number of steps
 - Task-driven: offline and online qualitative studies

[6] One click mining: Interactive local pattern discovery through implicit preference and performance learning. M. Boley, B. Kang, P. Tokmakov, M. Mampaey, S. Wrobel. IDEAS (ACM SIGKDD Workshop), 2013.

References

- [1] MRI: Meaningful Interpretations of Collaborative Ratings, S. Amer-Yahia, Mahashweta Das, Gautam Das and Cong Yu. In PVLDB 2011.
- [2] Exploring Rated Datasets with Rating Maps Sihem Amer-Yahia, Sofia Kleisarchaki, Naresh Kumar Kolloju, Laks V.S. Laskhmanan, Ruben H. Zamar WWW 2017
- [3] Multi-Objective Group Discovery on the Social Web. B. Omidvar-Tehrani, S. Amer-Yahia, P. F. Dutot, and D. Trystram. ECML/PKDD 2016.
- [4] Mahashweta Das, Saravanan Thirumuruganathan, Sihem Amer-Yahia, Gautam Das, Cong Yu. Who Tags What? An Analysis Framework. Proceedings of the VLDB Endowment (PVLDB), VLDB Endowment, 2012, 5 (11), pp.1567-1578.
- [5] Interactive User Group Analysis. B. Omidvar-Tehrani, S. Amer-Yahia, A. Termier. CIKM 2015.
- [6] One click mining: Interactive local pattern discovery through implicit preference and performance learning. M. Boley, B. Kang, P. Tokmakov, M. Mampaey, S. Wrobel. IDEAS (ACM SIGKDD Workshop), 2013.