PhD thesis

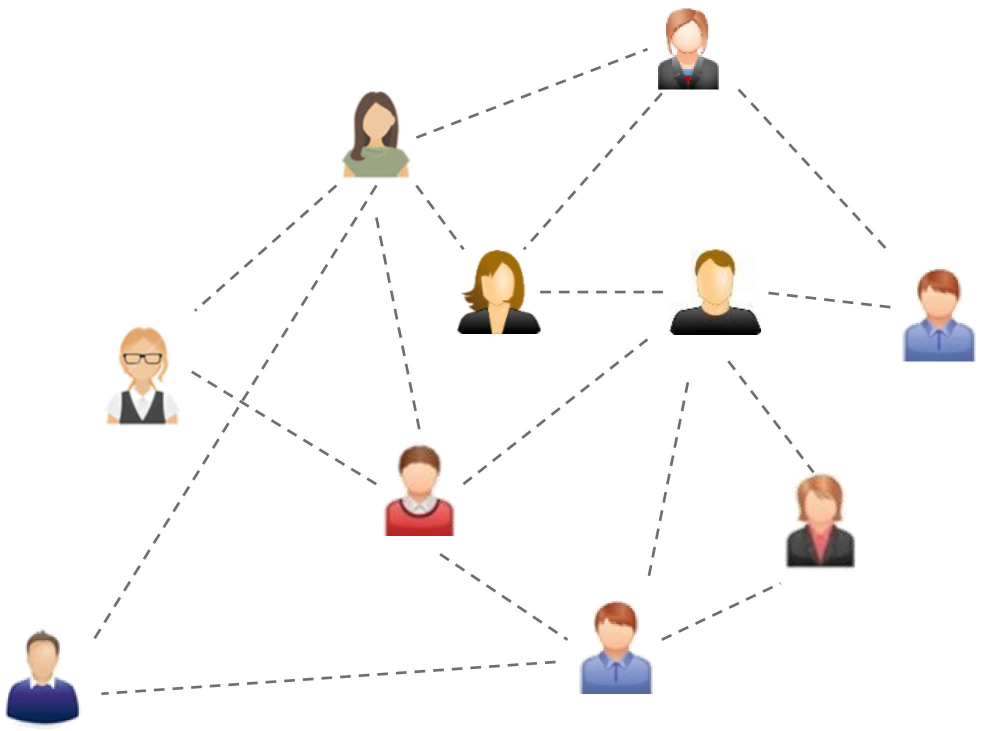
Contact recommendation in social networks: Algorithmic models, diversity and network evolution

Javier Sanz-Cruzado
under the supervision of
Pablo Castells

Online social networks

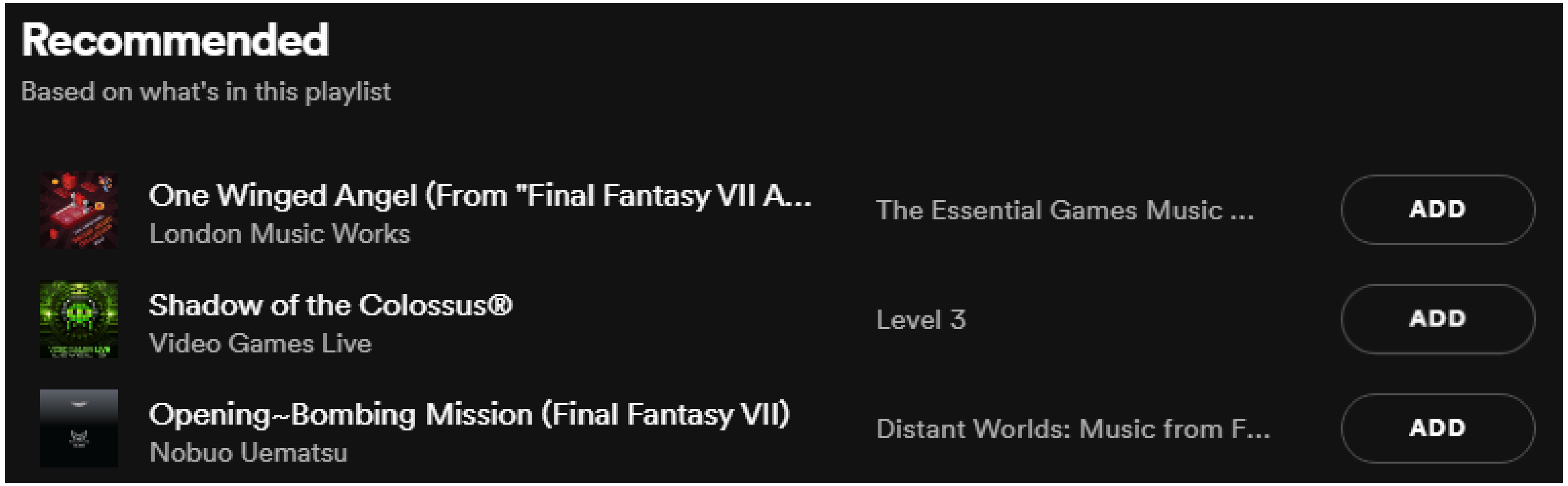
Online social networks (II)

- Establish new connections
- Communication
- Share and receive information
- Changes to our society
  - Politics
  - Privacy
  - Lifestyle
  - Communication



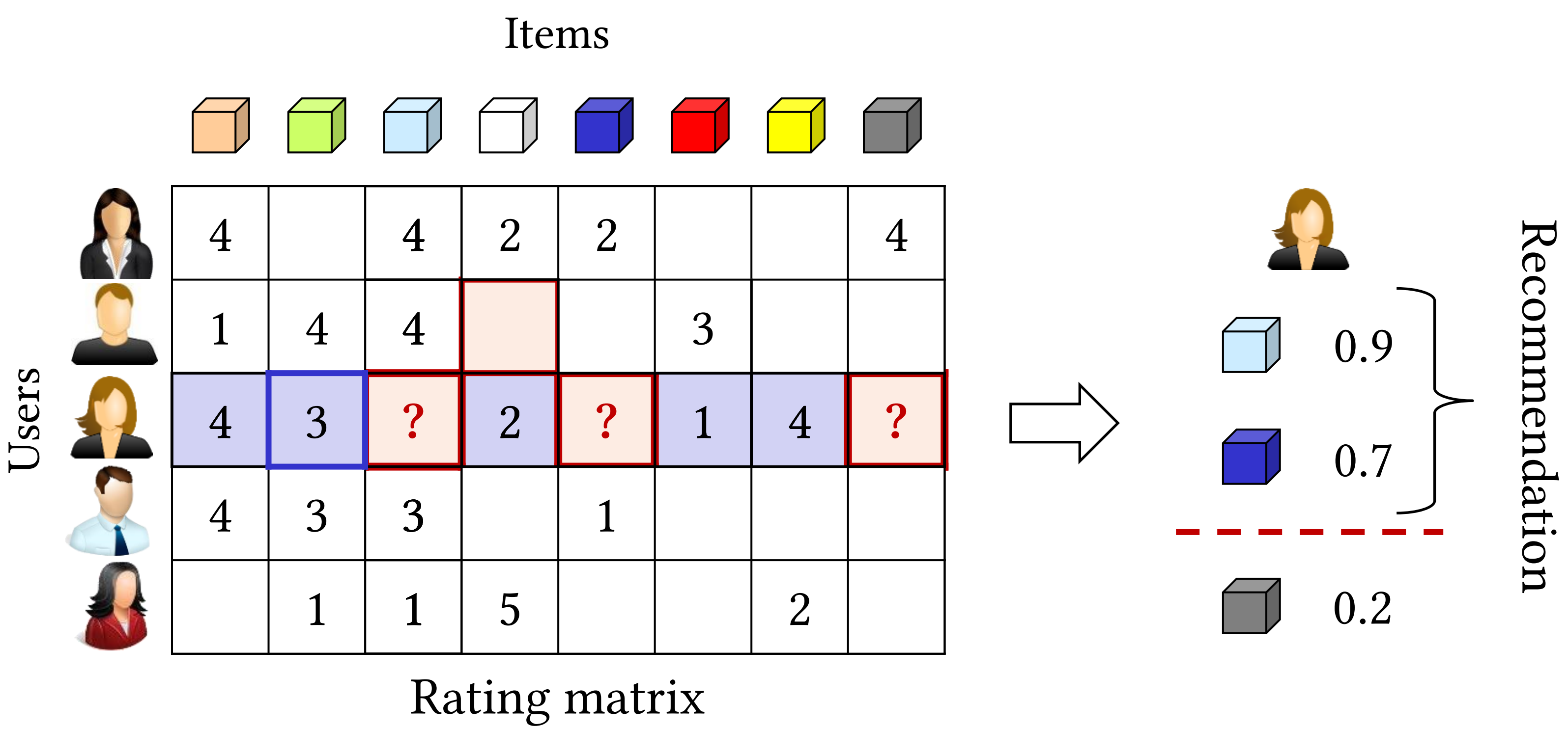
Recommender systems

- **Goal:** From past user interactions, suggest items they might be interested in.

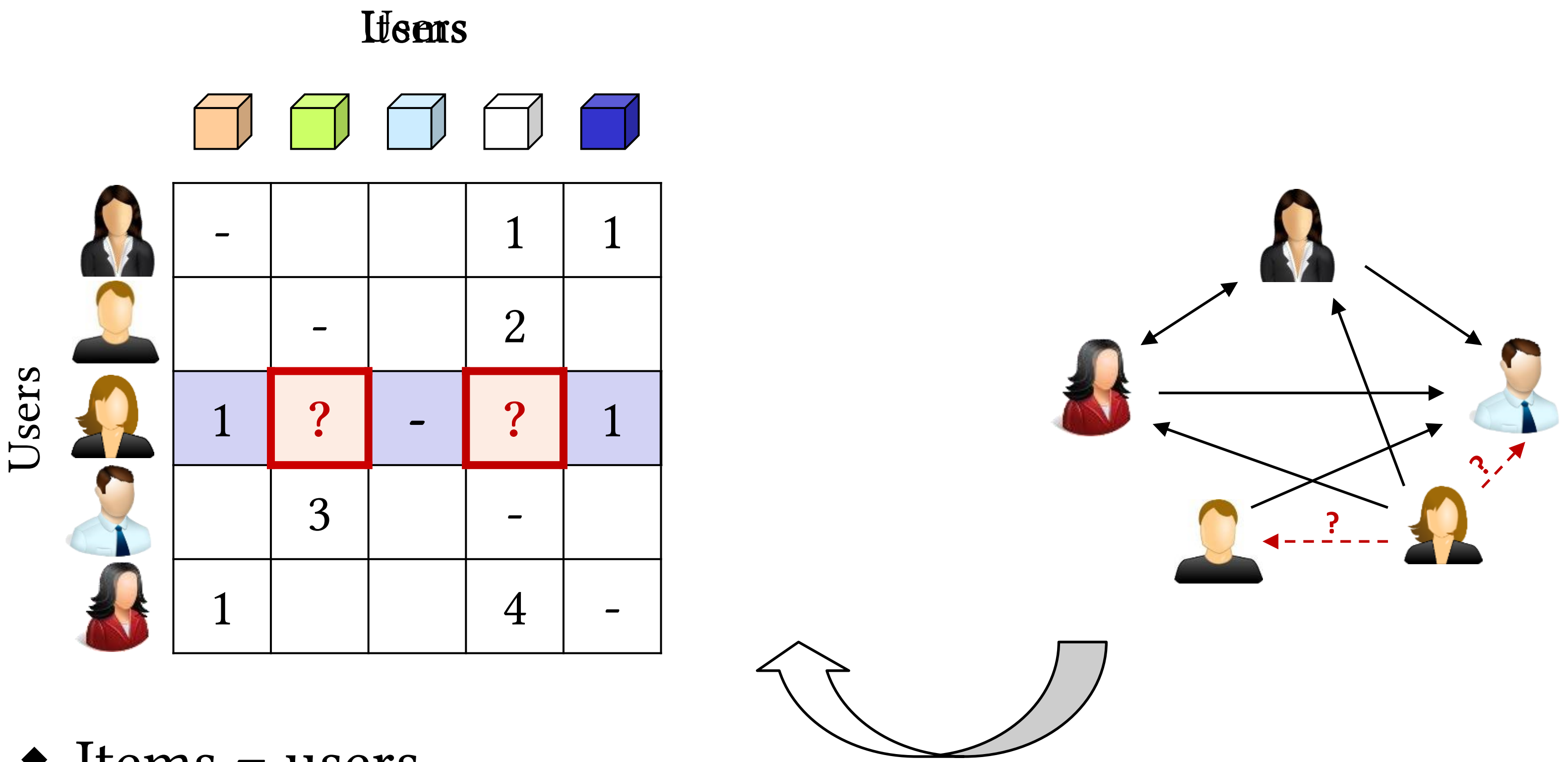


- Multiple domains
  - Audiovisual content: Netflix, Spotify
  - E-commerce: Amazon, eBay
  - Academic publications: Google Scholar, Mendeley

Recommendation task

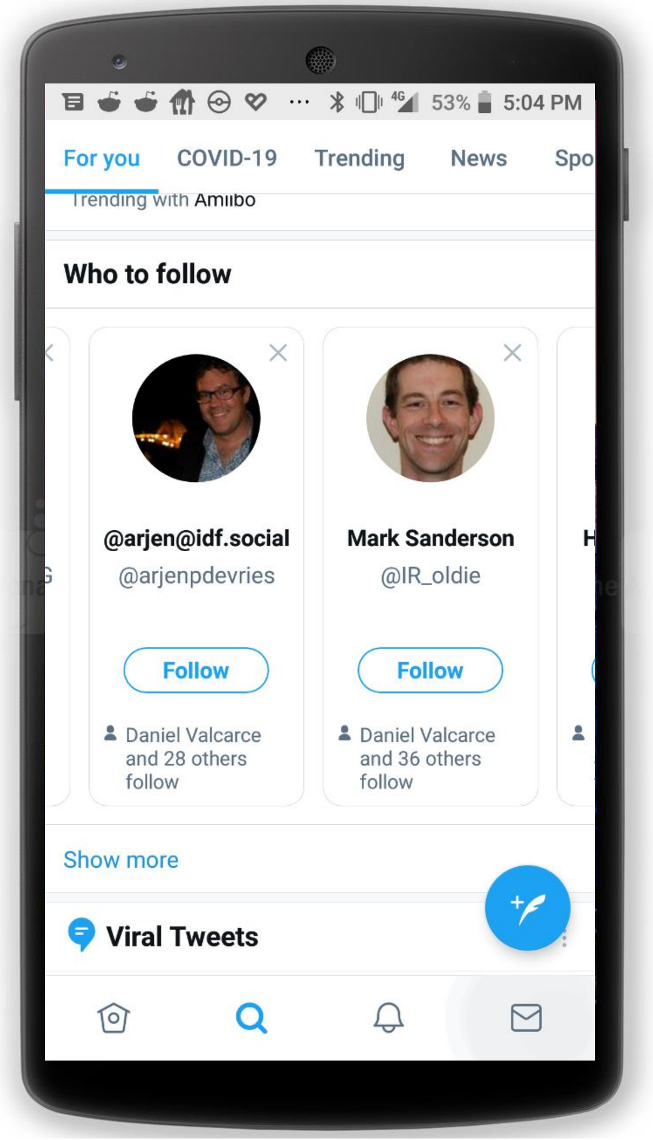


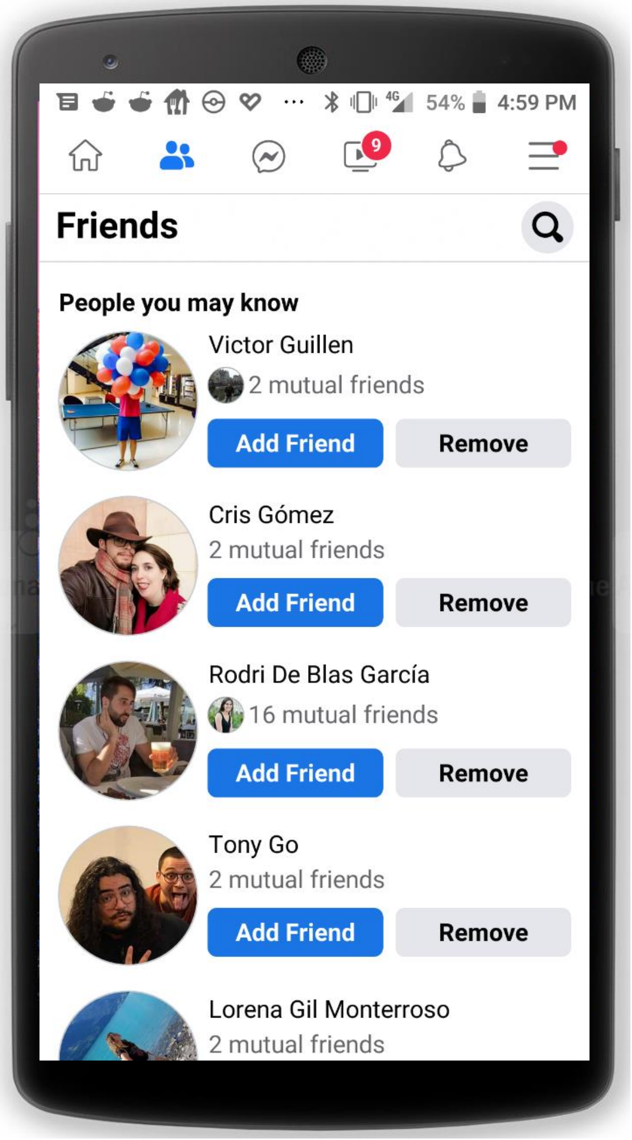
Contact recommendation



- Items = users
- Availability of social relationships
- Rating matrix = adjacency matrix

Contact recommendation examples





Why contact recommendation?

- Particular characteristics
  - Development of new methods
  - Use of social network analysis

- Creation of new links
  - Main asset of online social networks
  - Communication channels
  - Source of information
  - Increase engagement of users

Research goals

- **RG 1 : Algorithmic models**

  Explore the adaptations of text information retrieval (IR) models to the contact recommendation task.

Network evolution

- **RG2 : Diversity**

  Study the effect of contact recommendation on the properties of social networks.

- **RG3 : Recommendation cycle**

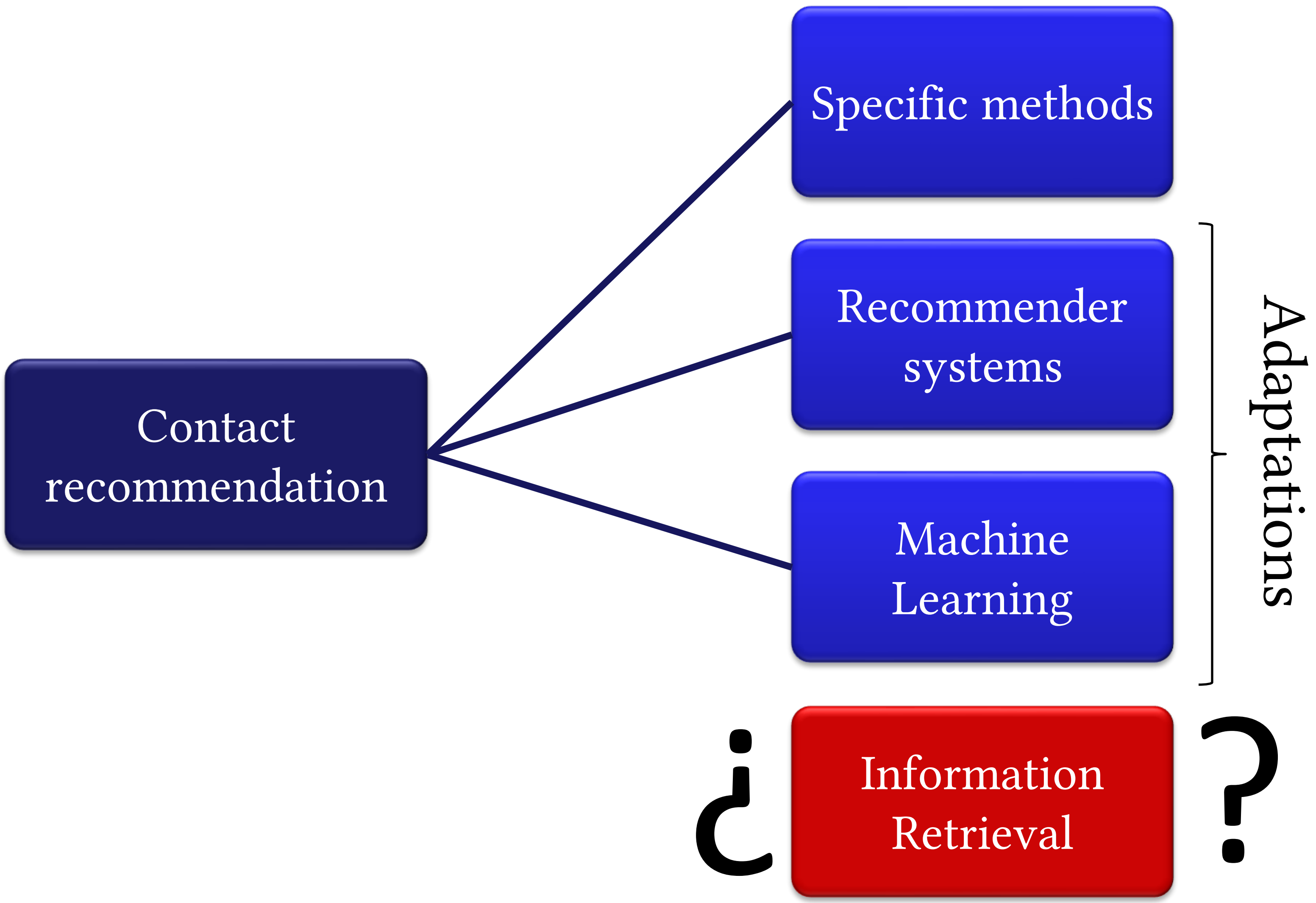
  Explore contact recommendation as a cyclic task, and develop interactive approaches to deal with it.

Outline

1. Adaptation of IR models
   - Relation between IR and contact recommendation
   - Advanced IR models
2. Beyond accuracy
   - Structural diversity
   - Effects on information diffusion
3. Interactive recommendation
4. Conclusions and future work

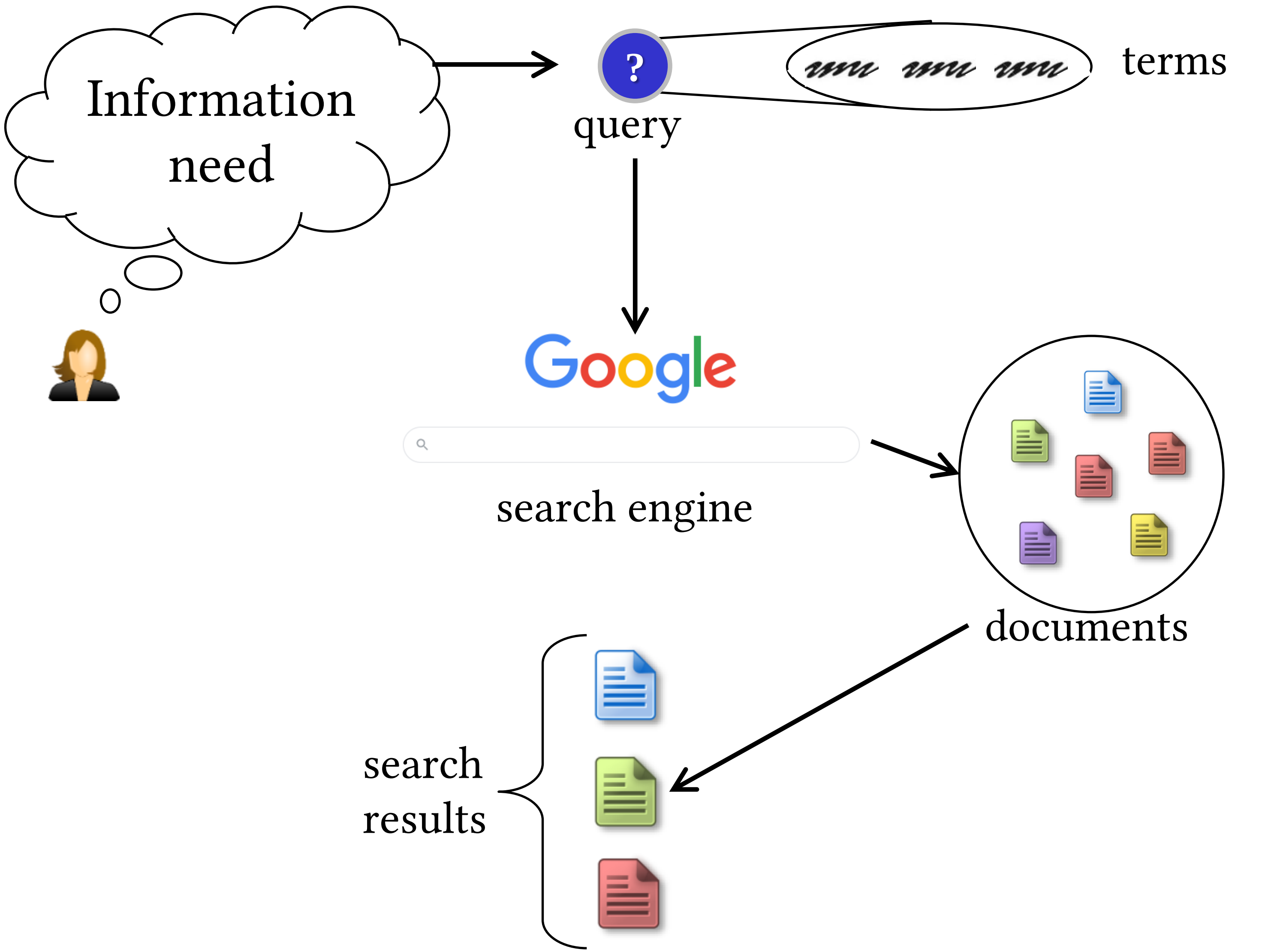
1 Adaptation of IR models

Motivation



Contact recommendation
Specific methods
Recommender systems
Machine Learning
Adaptations
¿ Information Retrieval ?

Text information retrieval

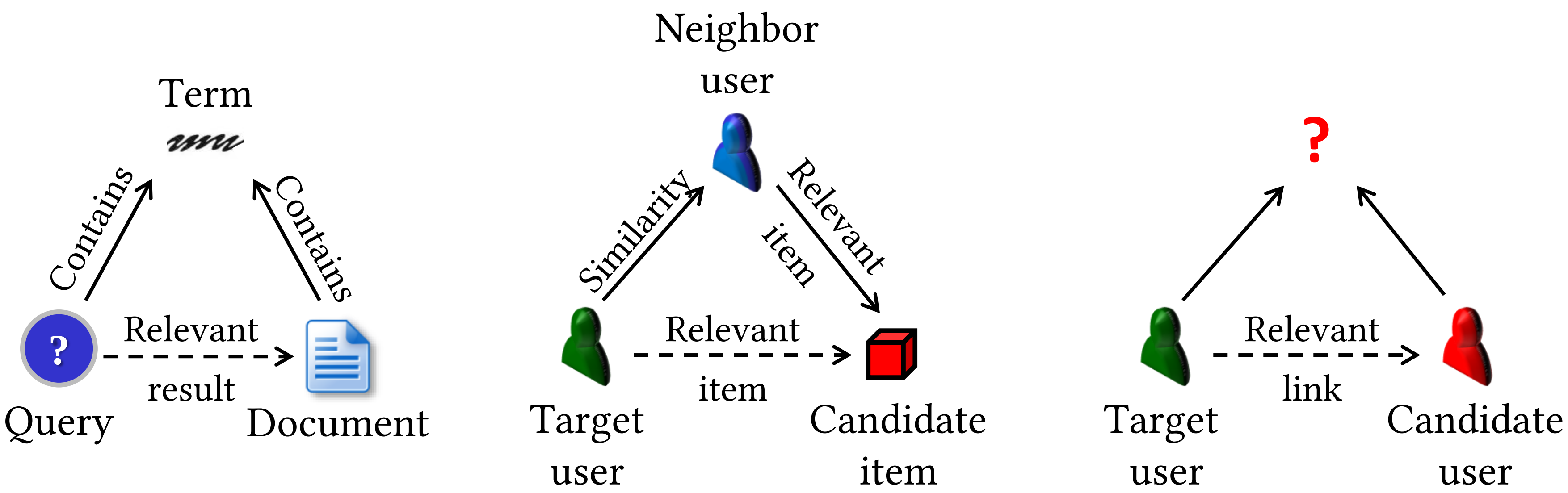


Relation between tasks

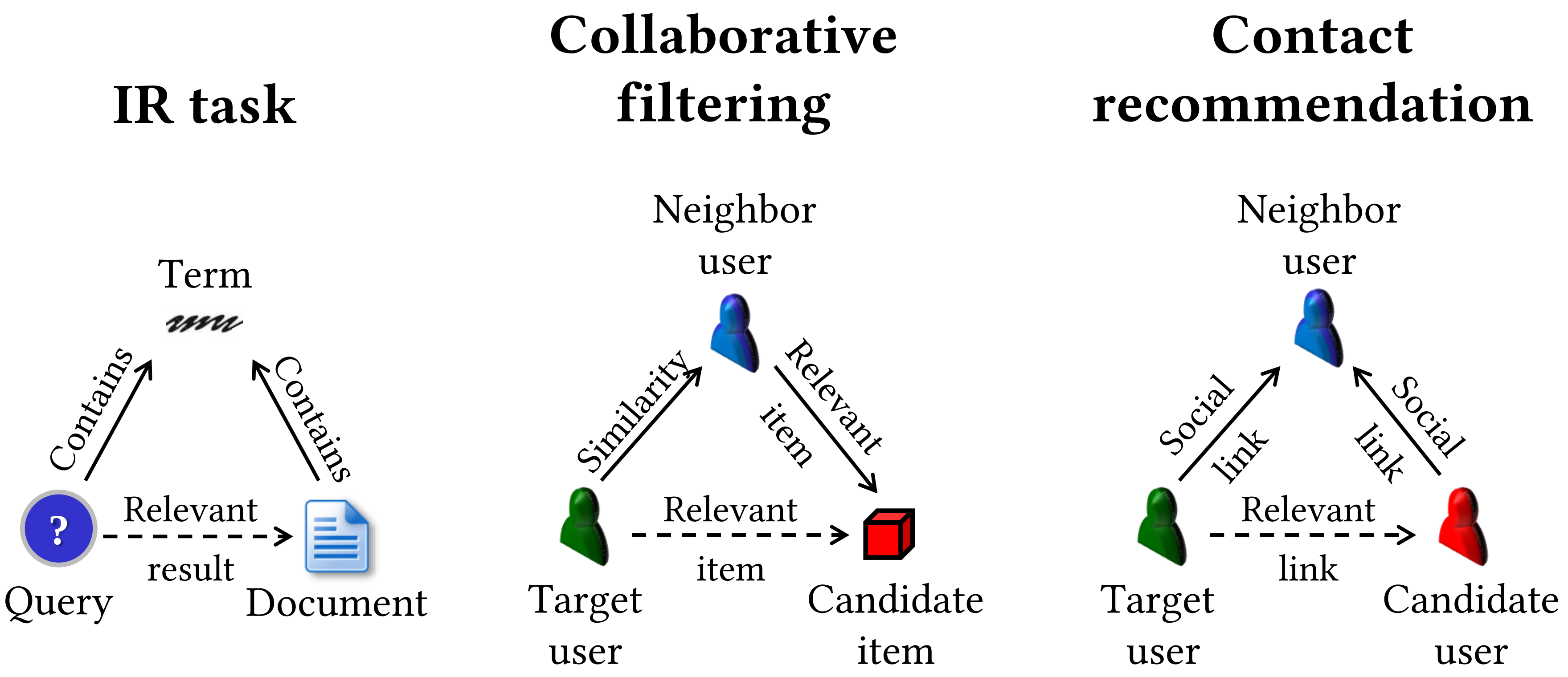
IR task

Collaborative filtering

Contact recommendation



Relation between tasks



An example: BM25

Text IR:

$$f_q(d) = \sum_{t \in d \cap q} \frac{(k+1)\,\mathrm{freq}(t,d)}{k\left(1 - b + \frac{b|d|}{\mathrm{avg}_{d'}|d'|}\right) + \mathrm{freq}(t,d)} \mathrm{RSJ}(t)$$

$$RSJ(w) = \log \frac{|D| - |D_t| - 0.5}{|D_t| - 0.5}$$

Where

- d: document → $\Gamma(v)$: candidate user
- q: query → $\Gamma(u)$: target user
- $t \in d \cap q$: term → $t \in \Gamma(u) \cap \Gamma(v)$: neighbor user
- D: set of all documents → $\mathcal{U}$: all users
- D_t: documents containing t → $\Gamma(t)$: v containing t in $\Gamma(v)$
- $\mathrm{freq}(t,d)$: frequency of t *in* d → $w(t,v)$: edge weight
- $|d|$: document d length → $\mathrm{len}(v) = \sum_{x \in \Gamma(v)} w(x,v)$

An example: BM25

Text IR:

$$f_q(d) = \sum_{t \in d \cap q} \frac{(k+1)\,\mathrm{freq}(t,d)}{k\left(1 - b + \frac{b|d|}{\mathrm{avg}_{d'}|d'|}\right) + \mathrm{freq}(t,d)} \mathrm{RSJ}(t)$$

$$\mathrm{RSJ}(w) = \log \frac{|D| - |D_t| - 0.5}{|D_t| - 0.5}$$

Contact recommendation:

$$f_u(v) = \sum_{t \in \Gamma(u) \cap \Gamma(v)} \frac{(k+1) w(t,v) \mathrm{RSJ}(t)}{k\left(1 - b + \frac{b \cdot \mathrm{len}(v)}{\mathrm{avg}_{v'}(\mathrm{len}(v'))}\right) + w(t,v)}$$

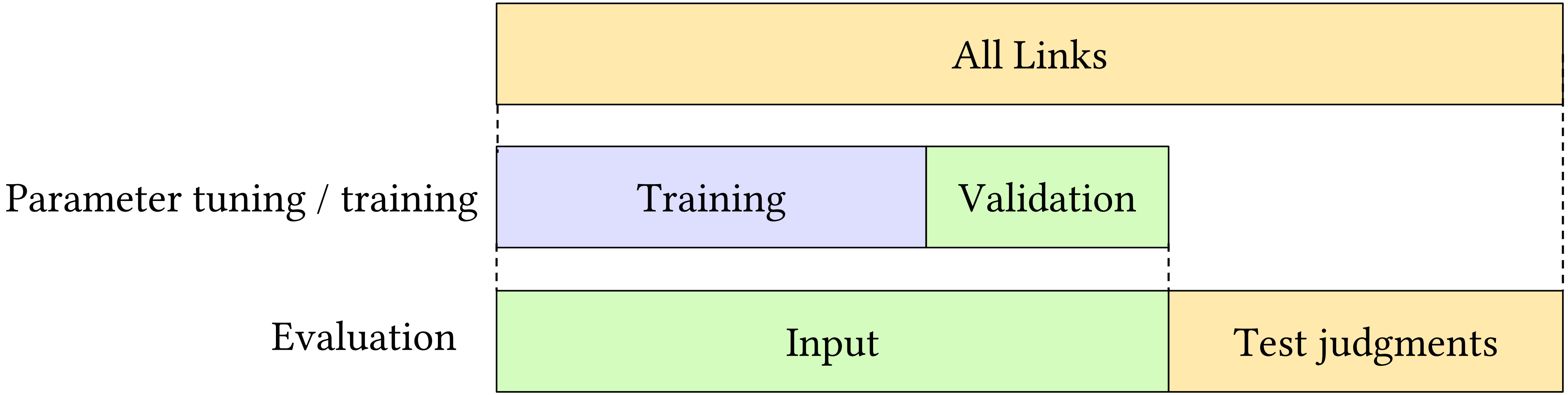
$$\mathrm{RSJ}(t) = \log \frac{|\mathcal{U}| - |\Gamma(t)| + 0.5}{|\Gamma(t)| + 0.5}$$

Experimental setup

- Offline evaluation
- Data from Twitter and Facebook
- **Twitter**
  - Snowball sampling
  - 2 samples
    - **1 month:** All tweets between 19th June and 19th July 2015
    - **200 tweets:** 200 last tweets by each user before 2nd August 2015
  - 2 graphs / dataset
    - Interaction networks: $(u, v) \in E$ if u mentions/retweets v
    - Follow networks
- **Facebook**
  - From Stanford Large Network Dataset Collection
  - Union of 10 ego-networks

Methodology

- Split:



- Hyperparameter selection: grid search (nDCG@10)
- Evaluate using IR metrics on test: nDCG@10, MAP@10

Dataset statistics

	Twitter 1-month		Twitter 200-tweets		Facebook
	Interactions	Follows	Interactions	Follows	
Users	9,528	9,770	9,985	9,964	4,039
Input edges	170,425	645,022	104,866	427,568	56,466
Test edges	54,335	81,110	21,598	98,519	17,643
Directed	✓	✓	✓	✓	✗
Weighted	✓	✗	✓	✗	✗
Split type	Temporal	Temporal	Temporal	Temporal	Random
Density	0.0018	0.0067	0.0013	0.0048	0.0087

Algorithms

- IR models:
  - **Probability ranking principle:** BM25, BIR, ExtremeBM25
  - **Language models:** Query likelihood (QLJM, QLD, QLL)
  - **Divergence from randomness:** PL2, DFRee, DFReeKLIM, DLH, DPH
  - Vector space model (VSM)
- General collaborative filtering
  - User-based / Item-based kNN (cosine similarity)
  - Implicit matrix factorization (iMF)
- Specific approaches
  - **Friends of friends:** Adamic-Adar, MCN, Jaccard, cosine similarity
  - **Random walks:** Personalized PageRank, Money,…
  - **Path-based:** Local Path Index, Katz…
- Sanity check: Random and most popular

Results (nDCG@10)

Algorithm	200-tweets Interaction	200-tweets Follows	Facebook
BM25	**<u>0.1097</u>**	0.1159	0.5731
BIR	0.1004	0.114	0.572
PL2	0.0983	0.1166	0.5712
VSM	0.0425	0.0787	0.5237
iMF	0.1035	**<u>0.1329</u>**	0.521
User-based kNN	0.0954	0.1297	0.5457
Item-based kNN	0.0724	0.1205	0.4542
Adamic-Adar	0.0997	0.114	0.5746
MCN	0.0948	0.111	0.5585
Resource allocation	0.0913	0.1117	**<u>0.5922</u>**
Personalized PageRank	0.063	0.0843	0.5891
Cosine	0.048	0.0768	0.4943
Popularity	0.0422	0.0397	0.0523
Random	0.0003	0.0018	0.003

- IR models are effective
  - BM25 among top 5
  - **Best:** 200-tweets interactions
  - VSM lowest performing IR model
- Rest of algorithms
  - Implicit MF is best
  - Adamic-Adar and MCN are competitive
  - Jaccard/cosine are not very competitive
  - Rest seem very graph dependent

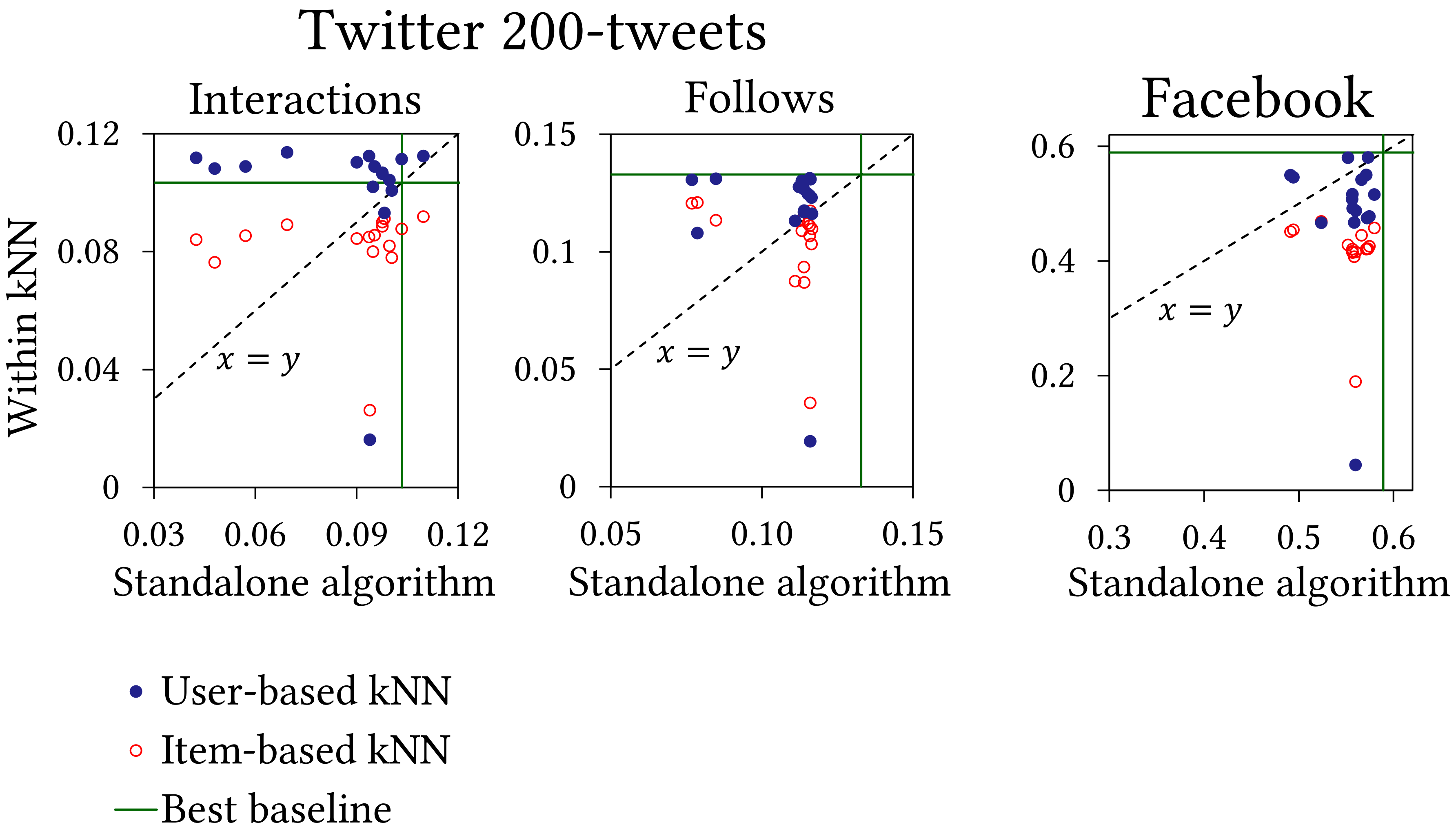
Can we do better?

	1-month		200-tweets		
Algorithm	Interaction	Follows	Interaction	Follows	Facebook
User-based kNN	0.1367	0.1413	0.0954	0.1297	0.5457
Item-based kNN	0.1174	0.1296	0.0724	0.1205	0.4542
Cosine	0.0393	0.0497	0.0480	0.0768	0.4943

User-based / Item-based kNN (cosine similarity) > Standalone cosine similarity

What if we try the same with IR models?

Results kNN + IR (nDCG@10)



Can we do even better?

- **Idea:** Learning to rank
  - Supervised machine learning models
  - Very effective in IR

- How does it work?
  1. Sample candidates
  2. Generate features for each target-candidate user pair
  3. Generate recommendation ranking

Our experiments

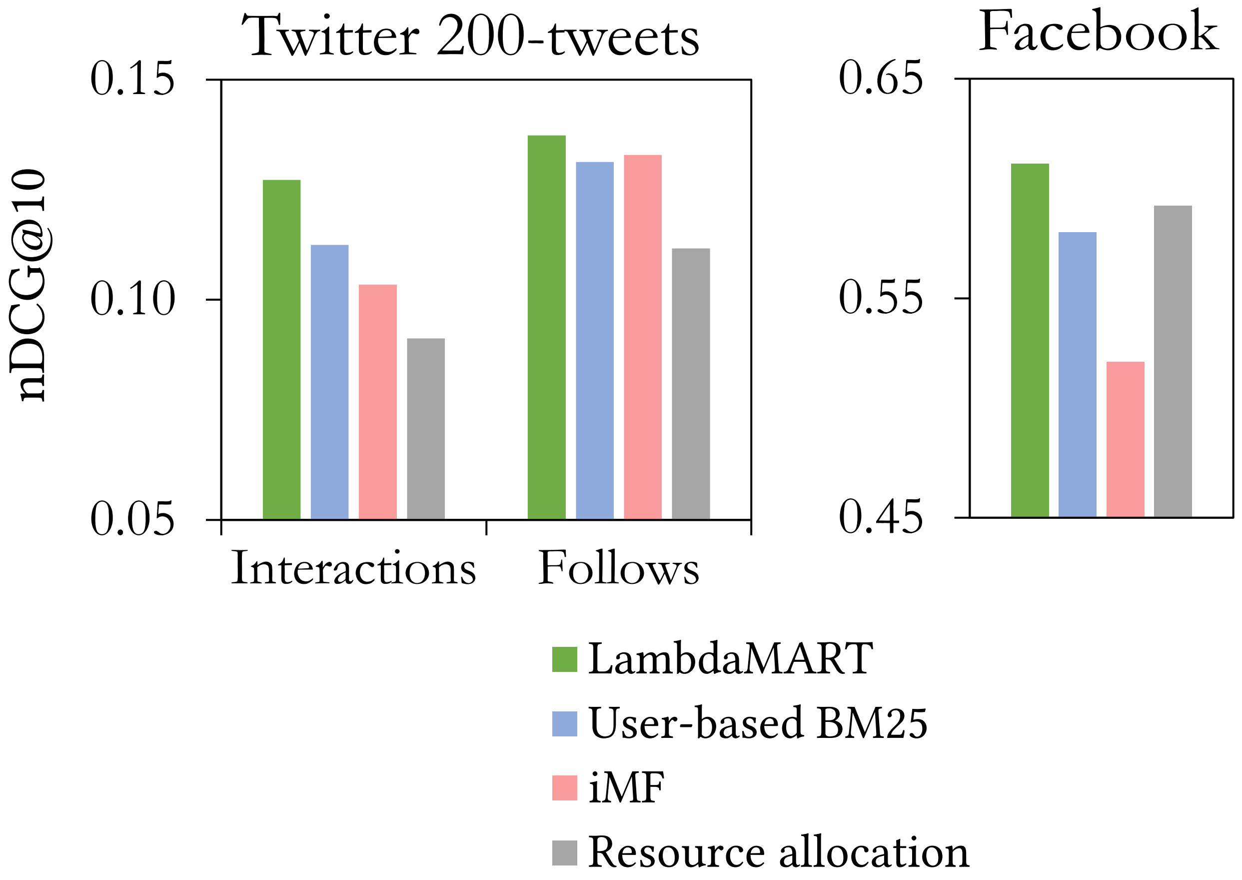
- **Learning to rank algorithm:** LambdaMART

- **Features:** Scores of contact recommendation methods
  - IR models
  - Friends of friends (FOAF) approaches
  - User-based / Item-based kNN + IR / FOAF

- **Sample suitable candidates:** use IR models

Learning to rank results

LambdaMART improves best recommendation baselines



Conclusions (RQ1)

- We can use IR models as contact recommendation algorithms
- Direct IR models are both effective and efficient (BM25)
- IR-based models are better as neighborhood selectors for kNN
- Learning to rank techniques improve the accuracy of best state of the art algorithms
- IR models are effective in three different roles in contact recommendation
  - Direct recommenders
  - Neighborhood selectors in kNN
  - Samplers and features in learning to rank

2 Beyond accuracy in contact recommendation

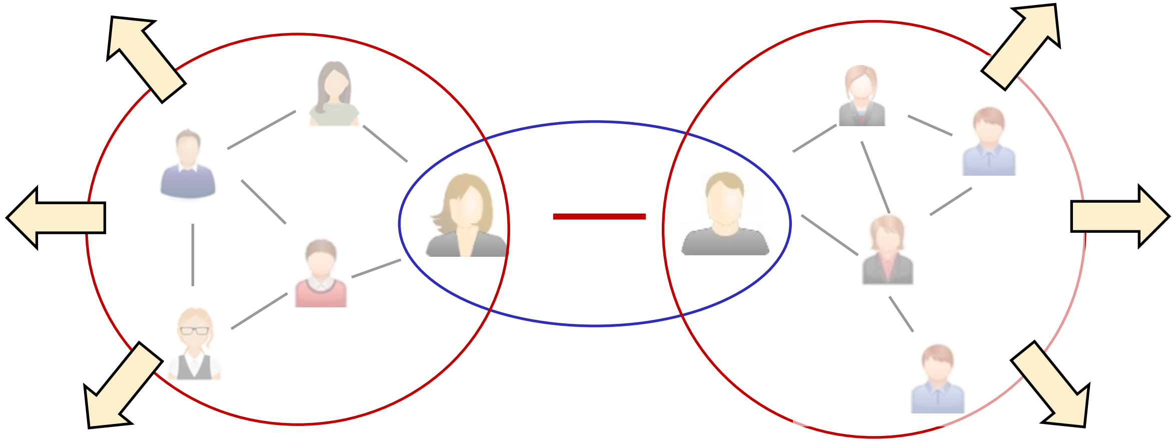
Accuracy

- Fundamental goal of contact recommendation

- Increase network density

- Limitations:
  - Local perspective: average over isolated users
  - Narrow perspective: one-dimensional utility

Beyond accuracy



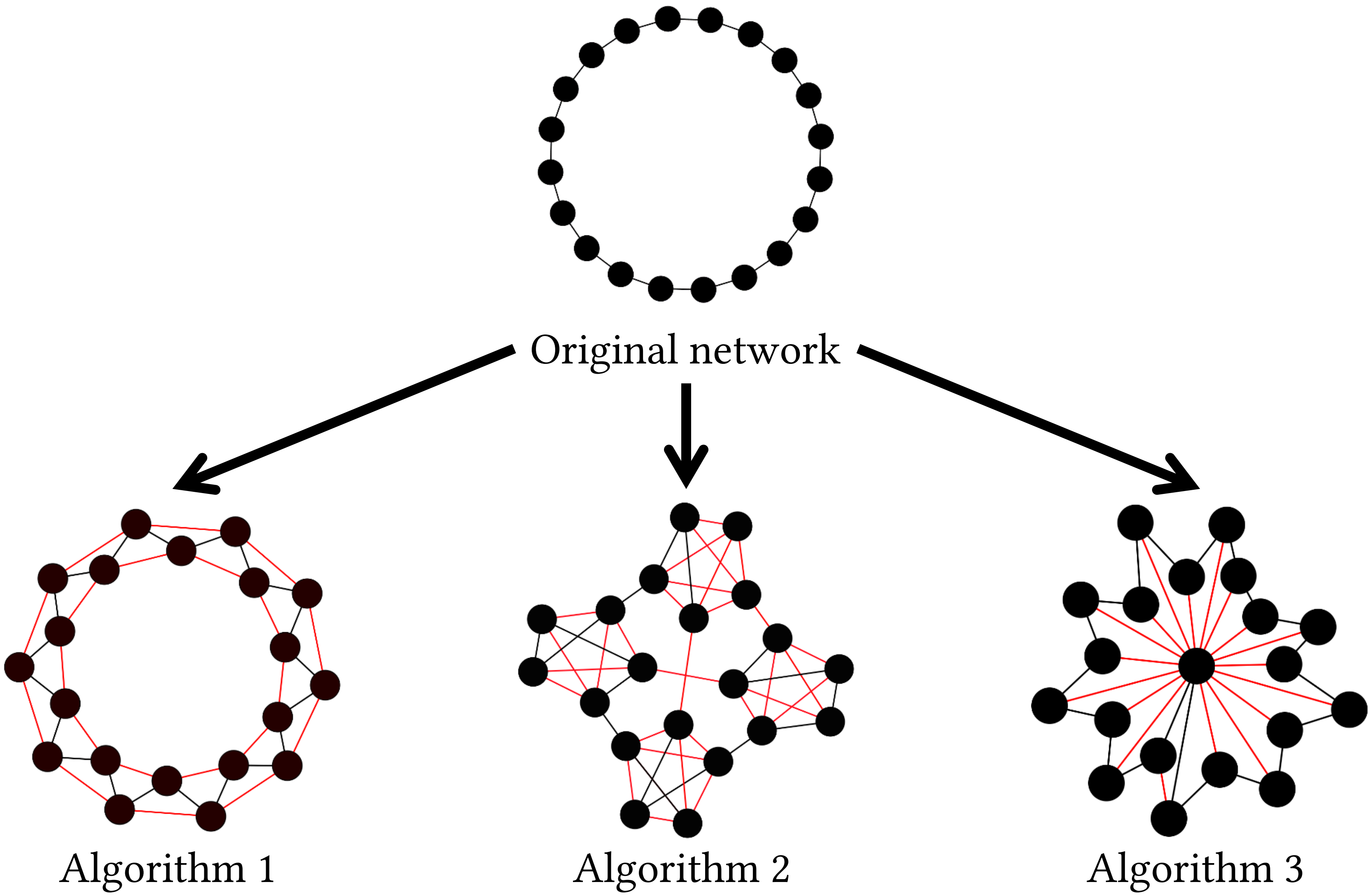
- Users in the network are not isolated
- A few links can cause global effects
- Different links – different effects
- Contact recommendation
  - 500 million new links/month on Twitter (as of 2015)
  - Potential to drive network evolution

Goals

1. Define suitable metrics to measure global benefits of recommendation

2. What do the metrics really mean? Do they capture relevant aspects of network functionality?

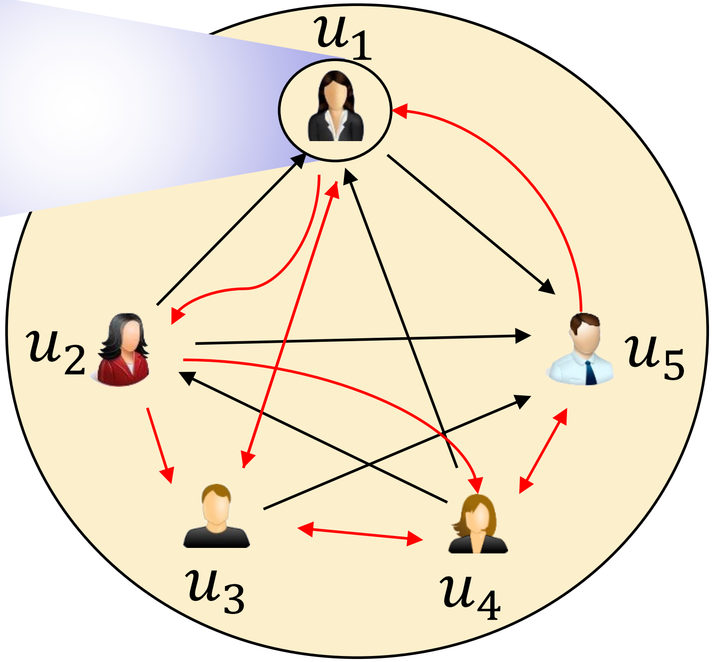
Effects on network structure



How to measure?

User	Score
u_2	0.9
u_3	0.8
u_4	0.1

Recommendation ranking



Structural metric

34

Potentially relevant structural features of social networks

- Structural diversity
  - Source of novel information
  - Enrichment of the information flow
  - Related to the notion of **weak tie** (Granovetter, 1978)
- Strength of a tie
  - Measures the involvement of users in the tie
  - **Strong ties:** family, close friends.
  - **Weak ties:** people you meet in conferences, shopkeepers.
- In the network structure: **non-redundant links**

Weak ties: local notions

- Consider the direct environment of the link

- **Triadic closure:** minimum unit of structural redundancy

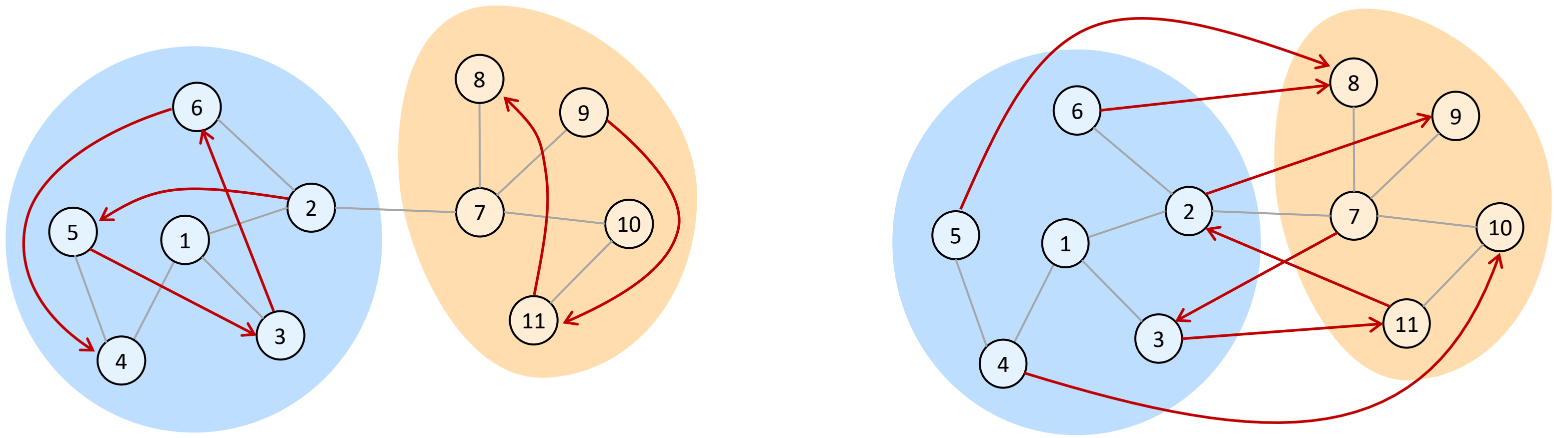
a) Non-redundant triad b) Redundant triad

- **Metric:** clustering coefficient complement

  Measures the proportion of non-redundant triads in the network

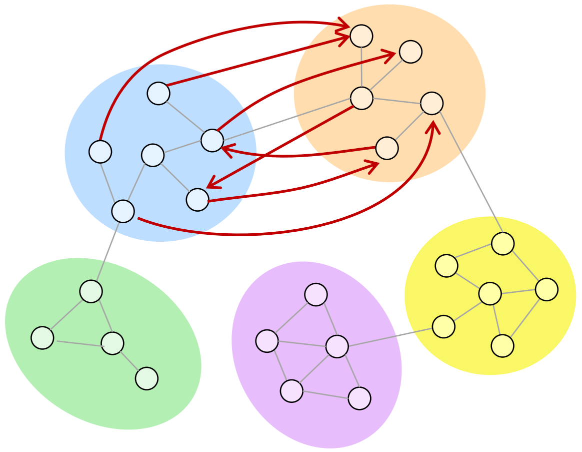
Weak ties: global notions

- **Weak ties:** links between communities
  - Tightly connected groups of nodes
  - Few connections outside the group
- **Modularity complement (MC):** number of weak ties

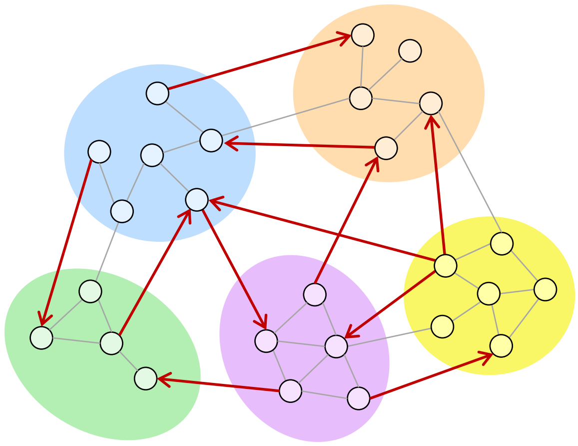


Low MC High MC

Weak ties: global notions (II)



Weak-link redundancy



Weak-link diversity

- **Community edge Gini complement (CEGC)**
  - **New metric**
  - Distribution of weak links between pairs of communities
  - Based on the Gini index

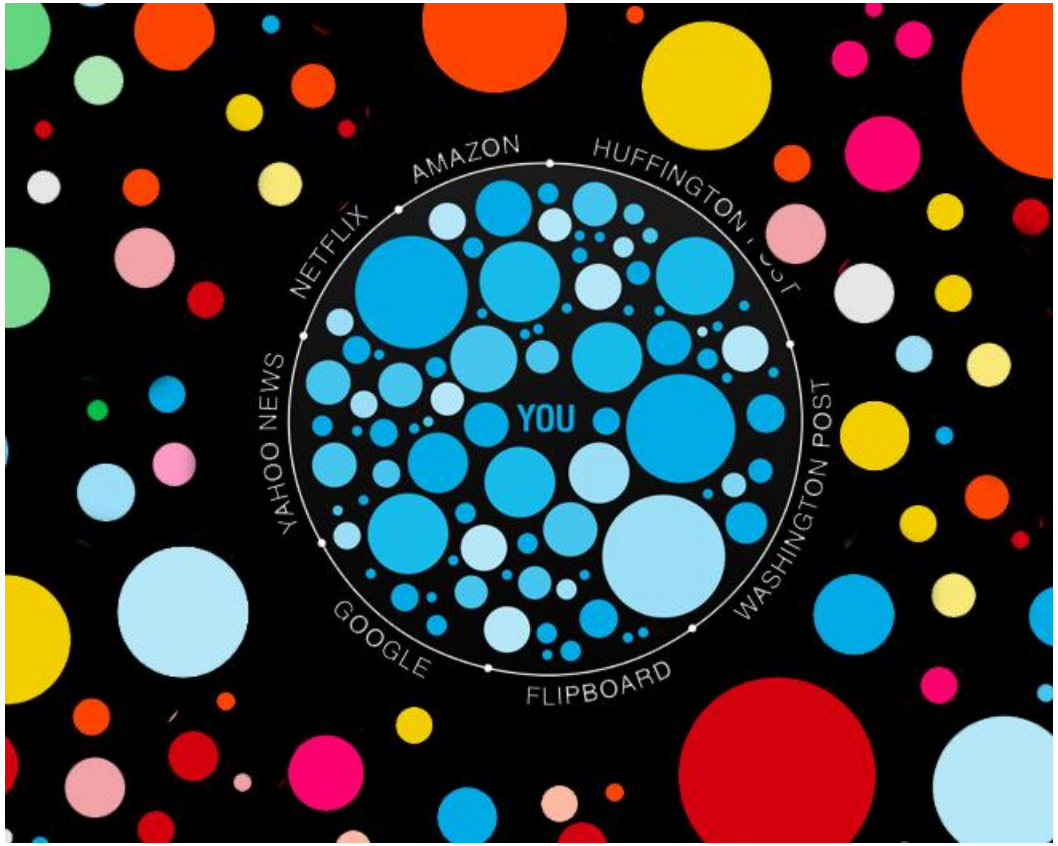
Effect of recommendation algorithms on structural diversity

Algorithm	nDCG@10	Clustering coefficient	Modularity	Community Gini
iMF	**0.139**	0.902	0.155	0.045
BM25	0.104	0.878	0.150	0.041
Adamic-Adar	0.098	0.882	0.149	0.041
MCN	0.092	0.879	0.145	0.040
Pers. PageRank	0.100	0.915	0.182	0.054
Popularity	0.057	0.924	**0.295**	0.061
Random	0.001	**0.952**	0.280	**0.091**
Original network	-	0.9437937	0.1463597	0.0390234

What do these numbers really mean for the network?

Filter bubbles

We analyze the potential of weak ties on reducing filter bubbles



Diffusion experiment

- Hypothesis

> The more structurally diverse the recommendation is,
> the more diverse and novel the information flow through
> the network will be.

- Experiment on interaction networks
  1. Start with a baseline: Implicit MF / BM25
  2. Apply gradual rerankers for optimizing a metric
  3. Extend the network with top k recommended links
  4. Run propagation of (real) tweets through the network
  5. Measure diffusion properties (novelty & diversity)

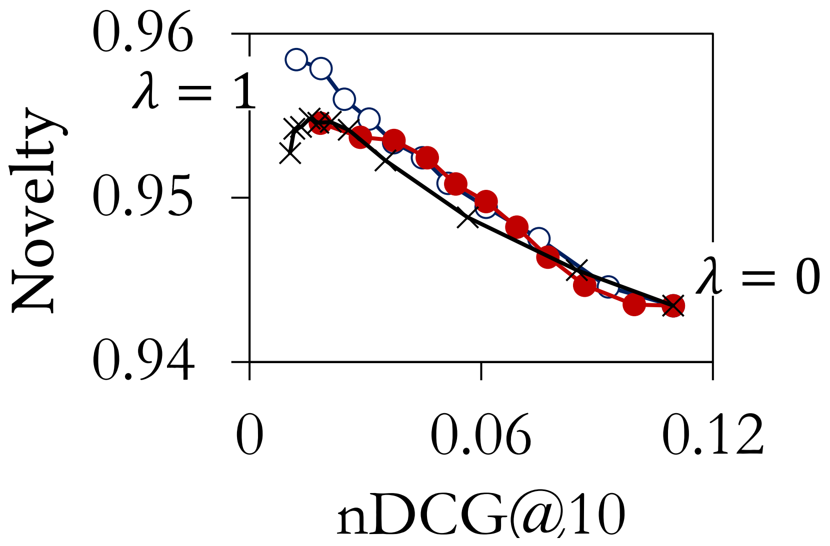
Diffusion properties

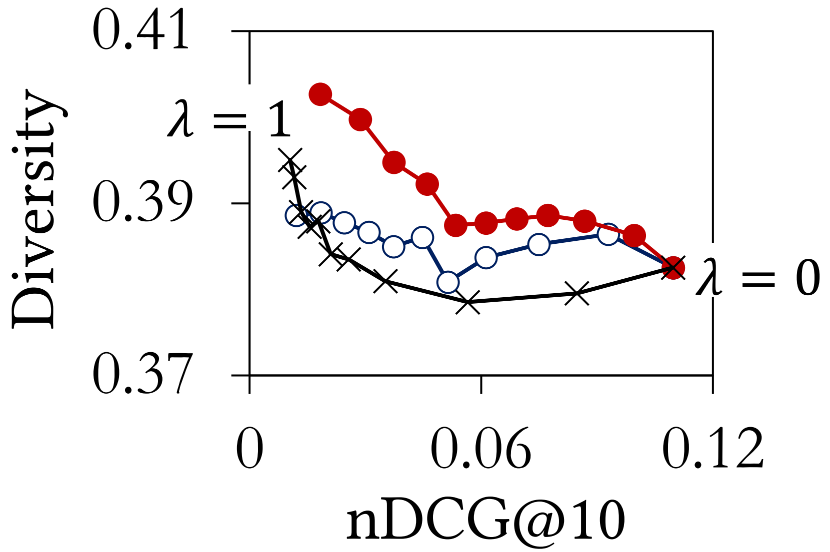
- Measured in terms of tweet hashtags (as topics)

- Novelty
  - Proportion of the hashtags unknown to the users.
  - Known hashtags: hashtags in their original tweets.

- Diversity
  - How evenly are hashtags propagated over the population
  - Complement of the Gini index

Results





Graph: Twitter 200-tweets interactions
Baseline: BM25

- Community Gini
- Modularity
- Clustering coefficient

Enhancing weak ties has positive effects in the novelty and diversity of the information flow

Conclusions (RG2)

- Accuracy is a partial perspective

- We propose evaluation perspectives beyond accuracy
  - Global network effects beyond (averaged) isolated user gains
  - New metrics elaborating on weak ties

- Enhancing the number of weak ties improves novelty & diversity of the information arriving to the users

3 Interactive recommendation

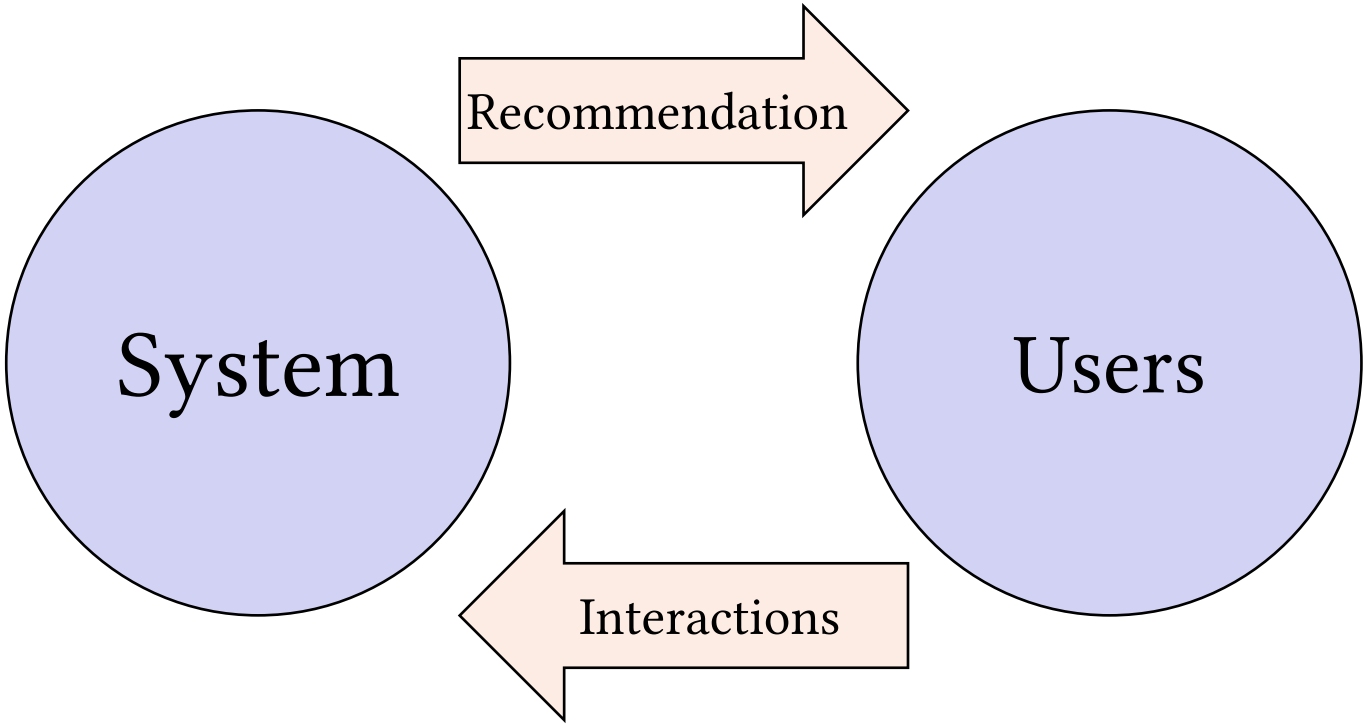
Motivation

- **Previous parts:** single recommendation step

- However
  - Recommendation does not work in a single step …
    but in an interactive process
  - Social networks are dynamic systems, constantly changing
  - And so recommender systems are

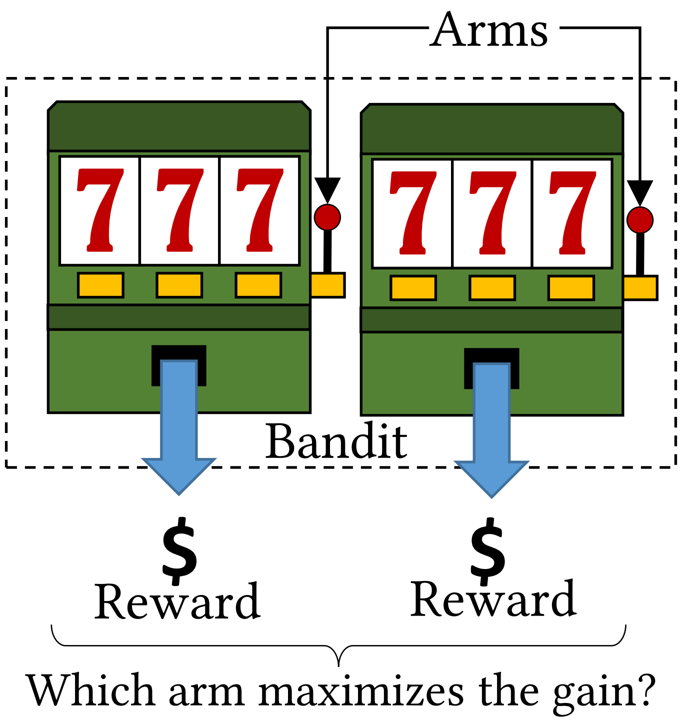
Interactive recommendation

- More realistic perspective
- Cyclic nature of recommendation



Multi-armed bandits

- Select the best among several actions (arms)
- Exploration vs. exploitation
  - Select arm with highest estimated value (exploit)
  - Select arm to gain knowledge (explore)



Bandit recommender systems

- Use bandits to generate recommendations
- **Personalized approaches:** contextual bandits
  - Change their actions depending on the context (user)
  - Examples:
    - Stochastic versions of collaborative filtering algorithms
    - Clusters of users / items (CLUB, COFIBA)
- **Relation between bandits and recommenders:**

Actions (arms) → Candidate items (users)

Rewards → Ratings

Context → Target user

Estimated arm value → Metric (e.g. CTR)

Our approach: nearest-neighbor bandit

- User-based kNN with stochastic neighborhood selection
- Uses a **Thompson sampling** bandit to select neighbors

- **Arms:** users in the system.
- **Estimated arm value:** conditional preference $p(u|w) = \frac{|\Gamma(u) \cap \Gamma(w)|}{|\Gamma(w)|}$

- How it works
  1. **Bandit:** Choose the optimal neighbor, w, for user u according to $p(u|w)$
  2. Neighbor w selects candidate user v according to $\mathrm{r}_{\mathrm{w}}(v)$
  3. Obtain the reward $r_u(v) \in \{0,1\}$
  4. Update $p(u|\hat{w})$ for all $\hat{w}$ s.t. $r_{\hat{w}}(v) > 0$

Thompson sampling

- **Assumption:** reward r follows a parametric distribution $p(r|\theta)$
  - Estimated arm value: $\mathbb{E}[r|\theta]$
  - Problem: θ unknown
- **Algorithm:** from previous data D
  1. Estimate $\hat{\theta}$ by sampling from $p(\theta|D)$
  2. Estimate the arm value as $\mathbb{E}[r|\hat{\theta}]$
- **Nearest neighbor bandit**
  - $p(u|v) \sim \text{Bernoulli}(p)$
  - $\theta = p \sim \text{Beta}\big(\alpha(u|v), \beta(u|v)\big)$

    $\alpha(u|v) = \alpha_0 + \#\text{Items both u and v like}$

    $\beta(u|v) = \beta_0 + \#\text{Items v likes, but u does not.}$

Extension: k neighbors

- Select k neighbors instead of one.
- Pick $\mathcal{N}_k(u)$: the k users maximizing the estimated $p(u|w)$

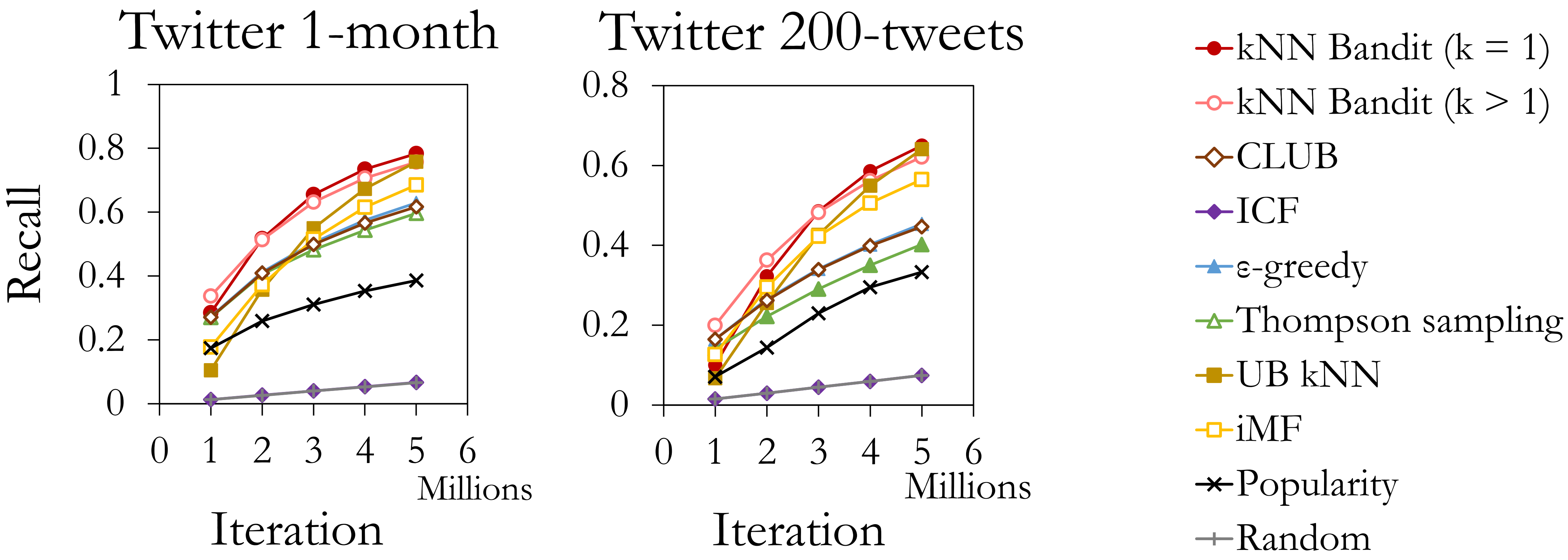
- Recommend the candidate user maximizing:

$$f_u(v) = \sum_{w \in \mathcal{N}_k(u)} p(u|w) r_w(v)$$

Experiments

- **Offline evaluation:** simulate feedback from offline data
- **Extreme cold start:** start with no ratings
- Random user selection (one at a time)
- **Metric:** cumulative recall
  - Fraction of discovered links at time t
  - Growth curve over time
- **Algorithms:**
  - Non-personalized bandits: ε-greedy, Thompson sampling
  - Personalized bandits: InterPMF, CLUB
  - Exploitation only: user-based kNN, iMF, most popular, random
  - Our approach ($k = 1, k > 1$)

Results (cumulative recall)



Conclusions (RG3)

- We have proposed a multi-armed bandit approach for interactive contact recommendation
  - Based on kNN
  - Uses a stochastic Thompson sampling strategy to select neighbors
- It provides relevant recommendations during the recommendation cycle.
- Our approach is more uncertainty-aware than myopic collaborative filtering approaches.

4 Conclusions

IRG
IR Group @UAM

UAM
UNIVERSIDAD AUTONOMA
DE MADRID

RG 1 : Algorithmic models

Explore the adaptations of text information retrieval (IR) models to the contact recommendation task.

Conclusions (RG1)

- We can use IR models as contact recommendation algorithms.
- IR models are both effective and efficient (BM25)
  - Direct recommenders (BM25)
  - Neighborhood selectors in kNN
  - Samplers and features in learning to rank
- IR-based models are better as neighborhood selectors for kNN
- Learning to rank techniques improve the accuracy of best state of the art algorithms.

RG 2 : Diversity

Study the effect of contact recommendation on the properties of social networks.

Conclusions (RG2)

- Accuracy is a partial perspective
- We consider evaluation perspectives beyond accuracy
  - Global network effects beyond (averaged) isolated user gains.
  - New metrics elaborating on weak ties.
- Enhancing the number of weak ties improves novelty & diversity of the information arriving to the users

RG 3 : Recommendation as a cycle

Understand contact recommendation as a cyclic task, and develop interactive approaches to deal with it.

Conclusions (RG3)

- We have proposed a multi-armed bandit approach for interactive contact recommendation
  - Based on kNN
  - Uses a stochastic strategy to select neighbors

- It improves medium to long-term accuracy

- Our approach is more uncertainty-aware than myopic collaborative filtering approaches.

Future work

- User studies and online evaluation
  - Complement our experiments
  - Determine the usefulness of our diversity dimensions
- Explore further relations with IR
  - Deep learning IR models
  - Other areas: query reformulation, relevance feedback
- Beyond accuracy
  - New dimensions: fairness
  - Find further benefits: reduce glass ceiling effect, radicalization
- Interactive recommendation
  - Explore other experimental settings
  - Analyze the evolution of the structural network properties

Thank you for your attention!

Questions?