

Sistemas Recomendadores Híbridos

IIC 3633 - Sistemas Recomendadores

Denis Parra

TOC

En esta clase

1. Motivación
2. Clasificación General
3. Modelos de Hibridización
4. Ejemplos

Motivación

Diferentes métodos tienen distintas debilidades y fortalezas

- Filtrado Colaborativo es preciso, pero sufre de sparsity, cold start y new item problem
- Filtrado Basado en contenido no sufre tanto por sparsity y permite con facilidad para extraer features del contenido. Sin embargo, también sufre de "new user problem", es menos preciso de el F.C. y presenta sobre-especialización.
- Knowledge-based: No los hemos visto hasta ahora. Casos típicos son Constraint-Based y Case-Based. Basados en un paradigma más interactivo, también los llaman "Conversacionales" (Burke, 2002). Su principal debilidad es el costo de mantener las reglas actualizadas.

Categorización de RecSys de Burke (2002)

Table I: Recommendation Techniques

Technique	Background	Input	Process
Collaborative	Ratings from U of items in I .	Ratings from u of items in I .	Identify users in U similar to u , and extrapolate from their ratings of i .
Content-based	Features of items in I	u 's ratings of items in I	Generate a classifier that fits u 's rating behavior and use it on i .
Demographic	Demographic information about U and their ratings of items in I .	Demographic information about u .	Identify users that are demographically similar to u , and extrapolate from their ratings of i .
Utility-based	Features of items in I .	A utility function over items in I that describes u 's preferences.	Apply the function to the items and determine i 's rank.
Knowledge-based	Features of items in I . Knowledge of how these items meet a user's needs.	A description of u 's needs or interests.	Infer a match between i and u 's need.

Ref: Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 12(4), 331-370.

Knowledge-Based Systems I

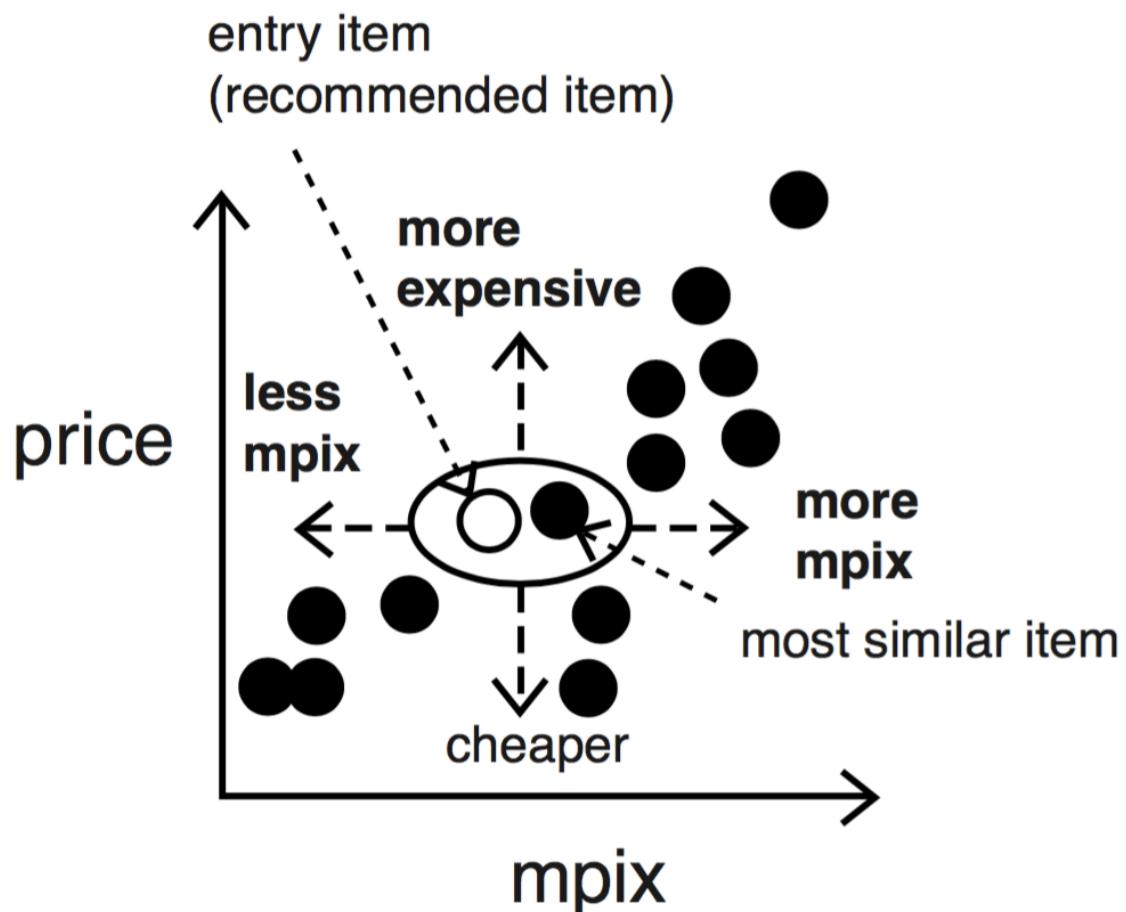


Figure 4.3. Critique-based navigation: items recommended to the user can be critiqued regarding different item properties (e.g., *price* or *mpix*).

Knowledge-Based Systems II



Fig. 7: The critiquing systems by Zhang et al. [2008], which visualizes compound critiques with icons and colors.

Ref: Jiyong Zhang, Nicolas Jones, and Pearl Pu. 2008. A visual interface for critiquing-based recom-mender systems. In Proceedings of the 9th Conference on Electronic Commerce (EC '08). 230–239.

Modelo Caja Negra de RecSys (Jannach et al. 2010)

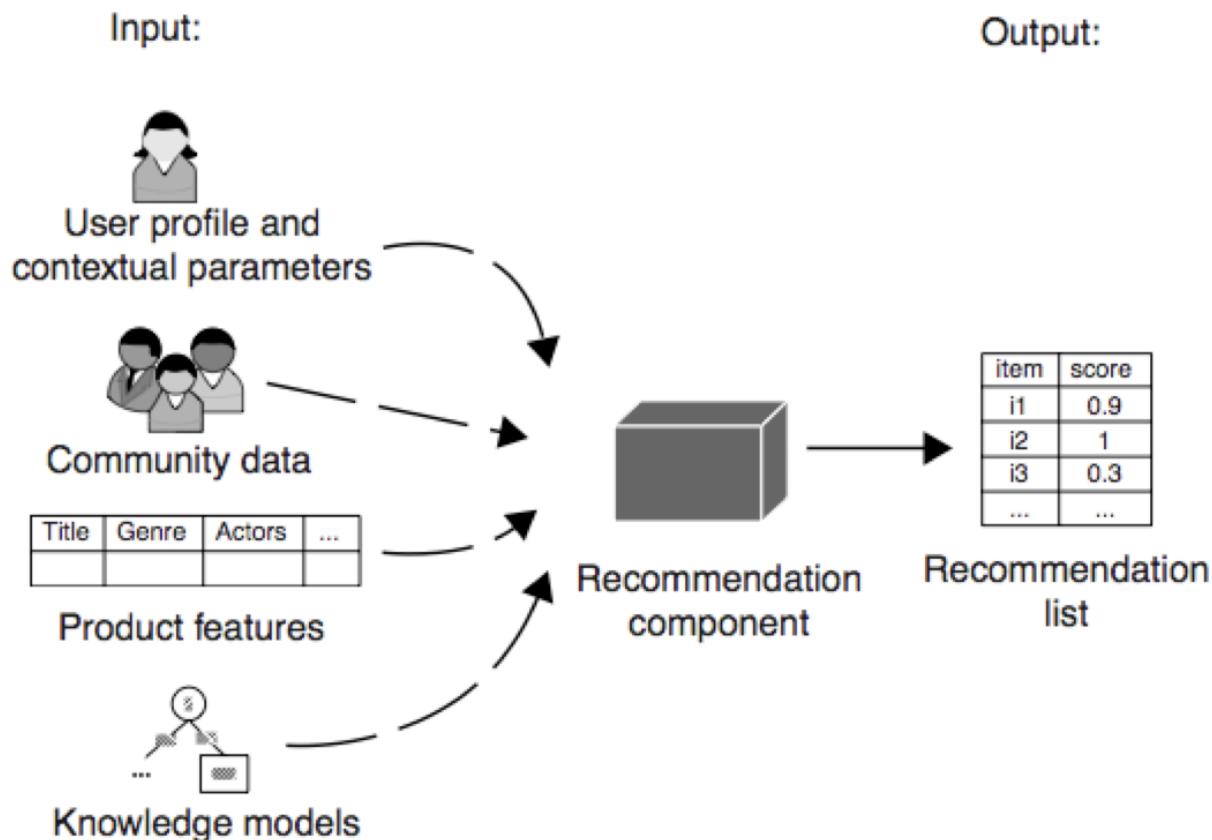


Figure 5.1. Recommender system as a black box.

Ref: Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). Recommender systems: 740 introduction. Cambridge University Press.

Combinar Métodos Content-based y Filtrado Colaborativo

Según (Adomavicius et al., 2005)

- Implementar métodos CF y CB separadamente y combinar las predicciones
- Incorporar características de métodos CB dentro de un método CF
- Incorporar características colaborativas dentro de modelo CB
- Construir un modelo que de manera unificada incorpore características basadas en contenido y colaborativas

Ref: Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6), 734-749.

Cómo combinar Métodos de Recomendación?

Burke (2002) distingue 7 estrategias de hibridización

Jannach (2012) resume las 7 estrategias en 3 diseños generales

- Monolítico
- Paralelizado
- Pipeline

7 Estrategias de Hibridización (Burke 2002)

Table III: Hybridization Methods

Hybridization method	Description
Weighted	The scores (or votes) of several recommendation techniques are combined together to produce a single recommendation.
Switching	The system switches between recommendation techniques depending on the current situation.
Mixed	Recommendations from several different recommenders are presented at the same time
Feature combination	Features from different recommendation data sources are thrown together into a single recommendation algorithm.
Cascade	One recommender refines the recommendations given by another.
Feature augmentation	Output from one technique is used as an input feature to another.
Meta-level	The model learned by one recommender is used as input to another.

Hibridización Monolítica

Table III: Hybridization Methods

Hybridization method	Description
Weighted	The scores (or votes) of several recommendation techniques are combined together to produce a single recommendation.
Switching	The system switches between recommendation techniques depending on the current situation.
Mixed	Recommendations from several different recommenders are presented at the same time.
Feature combination	Features from different recommenders are thrown together.
Cascade	One recommender feeds its output to another.
Feature augmentation	Features from one recommender are added to another.
Meta-level	The model learned by one recommender is used as input to another.

Hibridización Paralela

Table III: Hybridization Methods

Hybridization method	Description
Weighted	The scores (or votes) of several recommendation techniques are combined together to produce a single recommendation.
Switching	The recommendation techniques are switched at different times.
Mixed	Recommendations from several different recommenders are presented at the same time
Feature combination	Features from different recommendation data sources are thrown together into a single recommendation algorithm.
Cascade	One recommender refines the recommendations given by another.
Feature augmentation	Output from one technique is used as an input feature to another.
Meta-level	The model learned by one recommender is used as input to another.

Parallel

Hibridización Pipeline

Table III: Hybridization Methods

Hybridization method	Description
Weighted	The scores (or votes) of several recommendation techniques are combined together to produce a single recommendation.
Switching	The system switches between recommendation techniques depending on the current situation.
Mixed	Recommendations from several different recommenders are presented at the same time
Feature combination	Features from different recommendation data sources are thrown into a single recommendation.
Cascade	→ Pipeline → given by another.
Feature augmentation	Output from one technique is used as an input feature to another.
Meta-level	→ Pipeline → as input to

Hibridización Monolítica

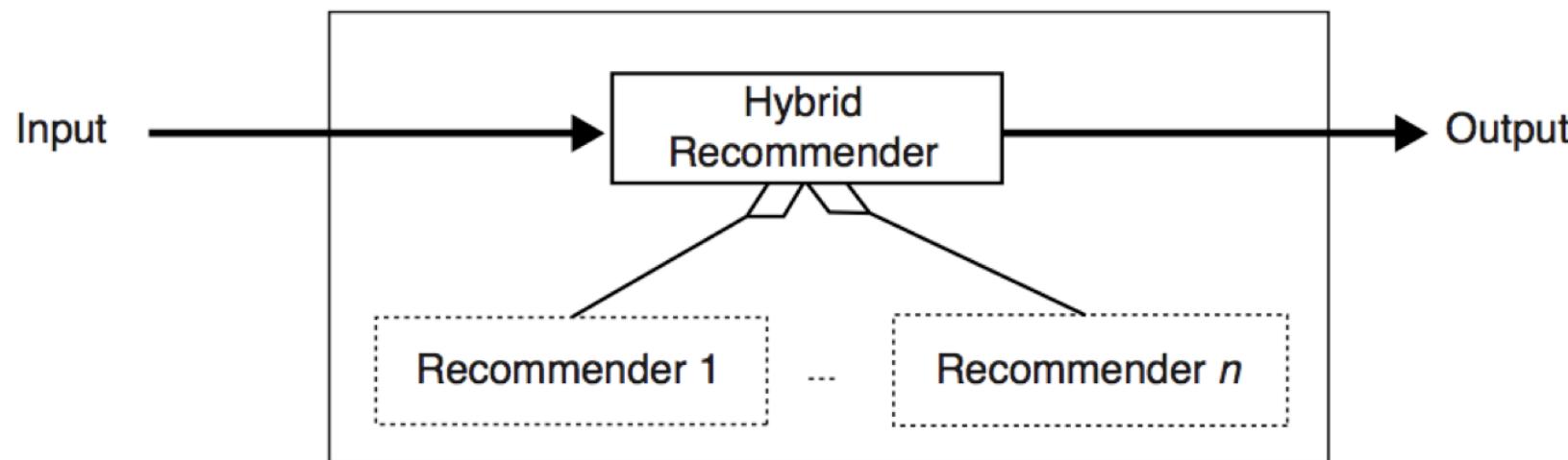


Figure 5.2. Monolithic hybridization design.

Estrategias de Combinación:

- Feature Combination
- Feature Augmentation

H. Monolítica: Feature Combination

Table 5.2. *Community and product knowledge.*

User	Item1	Item2	Item3	Item4	Item5	Item	Genre
Alice		1		1		Item1	romance
User1		1	1		1	Item2	mystery
User2	1	1			1	Item3	mystery
User3	1		1			Item4	mystery
User4					1	Item5	fiction

H. Monolítica: Feature Combination I

Table 5.3. *Hybrid input features.*

Feature	Alice	User1	User2	User3	User4
User likes many <i>mystery</i> books	true	true			
User likes some <i>mystery</i> books			true	true	
User likes many <i>romance</i> books					
User likes some <i>romance</i> books			true	true	
User likes many <i>fiction</i> books					
User likes some <i>fiction</i> books		true	true		true

H. Monolítica: Feature Combination II

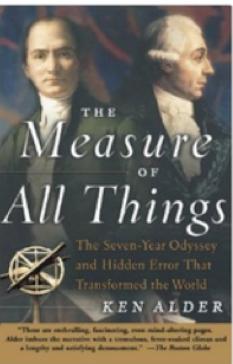
Table 5.4. *Different types of user feedback.*

User	R_{nav}	R_{view}	R_{ctx}	R_{buy}
Alice	n_3, n_4	i_5	k_5	\emptyset
User1	n_1, n_5	i_3, i_5	k_5	i_1
User2	n_3, n_4	i_3, i_5, i_7	\emptyset	i_3
User3	n_2, n_3, n_4	i_2, i_4, i_5	k_2, k_4	i_4

H. Monolítica: Feature Augmentation

- A diferencia de "feature combination""", este híbrido no sólo combina y procesa distintos tipos de entrada, sino que aplica algunos pasos de transformación más compleja.

[Look inside](#) ↴



The Measure of All Things: The Seven-Year Odyssey and Hidden Error That Transformed the World Paperback – October 1, 2003

by [Ken Alder](#) (Author)

★★★★★ 39 customer reviews

See all 11 formats and editions

Hardcover from \$0.01 **Paperback** \$20.74 

166 Used from \$0.01 44 Used from \$1.18
29 New from \$1.68 33 New from \$14.86
6 Collectible from \$9.85 3 Collectible from \$6.90

Customers Who Bought This Item Also Bought

 King Lear (Norton Critical Editions) by William Shakespeare ★★★★★ 5 Paperback \$13.95 	 A Philosophical Enquiry into the Origin of Our Ideas of the Beautiful by Edmund Burke ★★★★★ 10 Paperback \$10.30 	 Critique of Judgment by Immanuel Kant ★★★★★ 3 Paperback \$21.03 	 Sherlock Holmes: Selected Stories by Arthur Conan Doyle ★★★★★ 5 Paperback
--	---	--	--

H. Monolítica: Feature Augmentation



Usar estas “features” en un nuevo recomendador

Hibridización Paralela

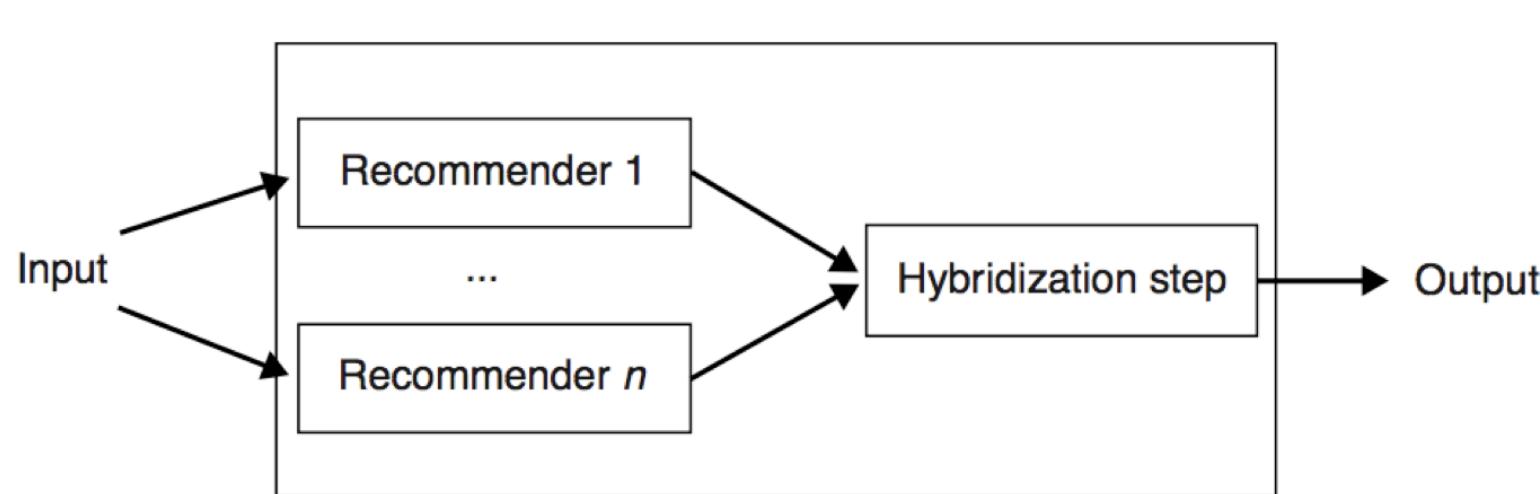


Figure 5.3. Parallelized hybridization design.

Tres mecanismos principales:

- Weighted
- Mixed
- Switching

H. Paralela: Weighted I

- La idea principal calcular una suma ponderada de los scores de items recomendados por diferentes recomendadores.

$$rec_{weighted}(u, i) = \sum_{k=1}^n \beta_k \times rec_k(u, i)$$

- La métrica de error MAE es ajustada:

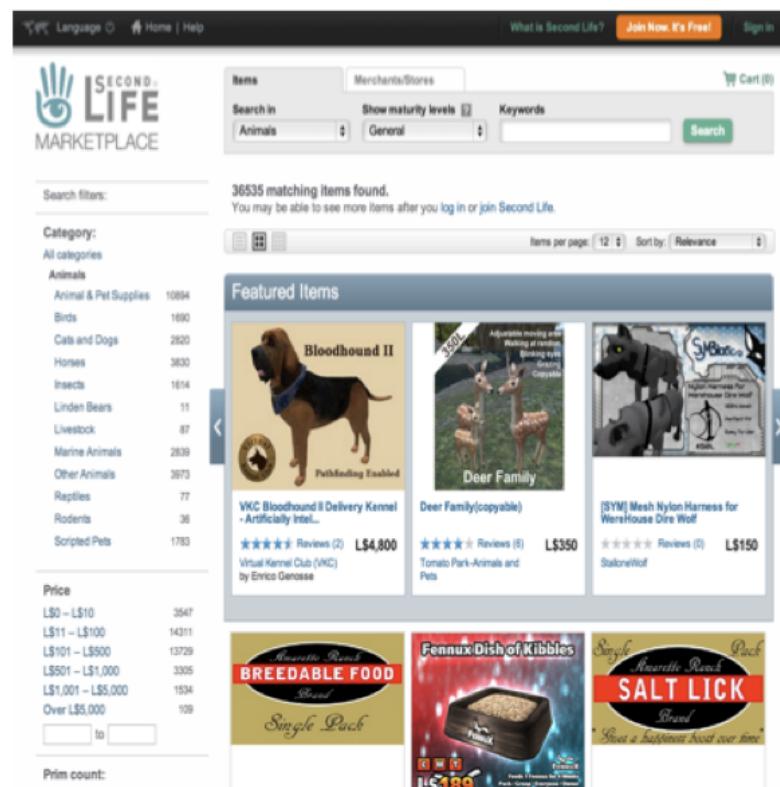
$$MAE = \frac{\sum_{r_i \in R} \sum_{k=1}^n \beta_k \times |rec_k(u, i) - r_i|}{|R|}$$

H. Paralela: Weighted II

$$W_{rec_i} = \sum_{s_j \in S} (W_{rec_i, s_j} \cdot W_{s_j})$$

rec_i	Item recomendado i
W_{rec_i}	Score combinado del item i
W_{rec_i, s_j}	Score del item I por el recomendador S_j
W_{s_j}	Peso del recomendador S_j

H. Paralela: Weighted III



(a) SecondLife store

The screenshot shows a user profile for 'Sylvia Tamalyn' (sylvia.tamalyn). The profile includes a photo of a woman with blonde hair, a 'Add Friend' button, a 'Follow' button, a 'Message' button, and a gear icon. Below the profile are tabs for Feed, About, Picks, and Groups. The feed shows a post from 'dresden.ceriano' with the caption: "Are you wearing your fur coat again, Tem? :p" about 4 hours ago. Another post by 'sylvia.tamalyn' shows a screenshot of her character in a field, with the caption: "Might as well play with WL while I am here :D" and a location pin for Shermerville NW Visit location. This post has 3 comments and 20 loves. A comment from 'empressazy' says: "love the imagery. man i am still workin with WL." and another from 'sylvia.tamalyn' says: "Thank you :)" both posted about 11 hours ago. A text input field at the bottom says "Write a comment...".

(b) SecondLife social stream

Fig. 1. Examples for a store in the marketplace and a social stream of an user in the online social network of the virtual world SecondLife.

H. Paralela: Weighted IV

- Ejemplo

Table 5.6. *Recommendations of weighted hybrid.*

item	rec_1 score	rec_1 rank	rec_2 score	rec_2 rank	rec_w score	rec_w rank
Item1	0.5	1	0.8	2	0.65	1
Item2	0		0.9	1	0.45	2
Item3	0.3	2	0.4	3	0.35	3
Item4	0.1	3	0		0.05	
Item5	0		0		0	

H. Paralela: Weighted V

- Asumiendo que Alice compró ítems 1 y 4

Table 5.7. *Dynamic weighting parameters, absolute errors, and MAEs for user Alice.*

β_1	β_2	item	r_i	rec_1	rec_2	error	MAE
0.1	0.9	Item1	1	0.5	0.8	0.23	0.61
		Item4	1	0.1	0	0.99	
0.3	0.7	Item1	1	0.5	0.8	0.29	0.63
		Item4	1	0.1	0	0.97	
0.5	0.5	Item1	1	0.5	0.8	0.35	0.65
		Item4	1	0.1	0	0.95	
0.7	0.3	Item1	1	0.5	0.8	0.41	0.67
		Item4	1	0.1	0	0.93	
0.9	0.1	Item1	1	0.5	0.8	0.47	0.69
		Item4	1	0.1	0	0.91	

H. Paralela: Mixed I

$$src(rec_i) = \left[\sum_{m_j \in M} \frac{1}{rank_{rec_i, m_j}} \times W_{m_j} \right] \times |M_{rec_i}|$$

M : The set of all methods available to fuse

$rank_{rec_i, m_j}$: rank–position in the list of a recommended item

rec_i : recommended item i

m_j : recommendation method j

W_{m_j} : weight given by the user to the method m_j using the controllable interface

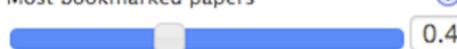
$|M_{rec_i}|$ represents the number of methods by which item rec_i was recommended

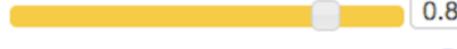
Slider

weight

H. Paralela: Mixed II

Tune weights of the recommender methods: (b)

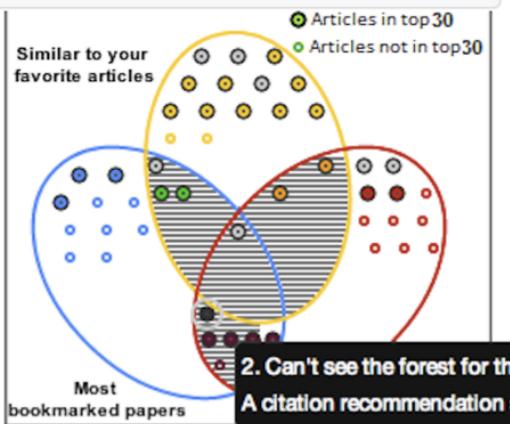
Most bookmarked papers  0.4

Similar to your favorite articles  0.8

Frequently cited authors in ACM DL  0.4

[Update Recommendation List →](#)

* Hover over circles to explore articles
* Click on the diagram to highlight subsets



Similar to your favorite articles

Most bookmarked papers

Articles in top 30

2. Can't see the forest for the trees?
A citation recommendation system

(a)

2. Can't see the forest for the trees? A citation recommendation system  [see abstract]

by C. Lee Giles, Cornelia Caragea, Adrian Silvescu, Prasenjit Mitra

3. When thumbnails are and are not enough: Factors behind users  [see abstract]

by Dan Albertson

7. Gendered Artifacts and User Agency  [see abstract]

by Andrea R. Marshall, Jennifer A. Rode

8. Two Paths to Motivation through Game Design Elements: Reward-Based Gamification and Meaningful Gamification  [see abstract]

by Scott Nicholson

9. Automatic Identifying Search Tactic in Individual Information Seeking: A Hidden Markov Model Approach  [see abstract]

by Zhen Yue, Shuguang Han, Daqing He

11. Old Maps and Open Data Networks  [see abstract]

by Werner Robitzka, Carl Lagoze, Bernhard Haslhofer, Keith Newman, Amanda Stefanik

14. Effects of User Identity Information On Key Answer Outcomes in Social Q&A  [see abstract]

by Erik Choi, Craig Scott, Chirag Shah

15. Ebooks and cross generational perceived privacy issues  [see abstract]

by Jennifer Sue Thiele, Renee Kapusniak

16. Toward a mesoscopic analysis of the temporal evolution of scientific collaboration networks  [see abstract]

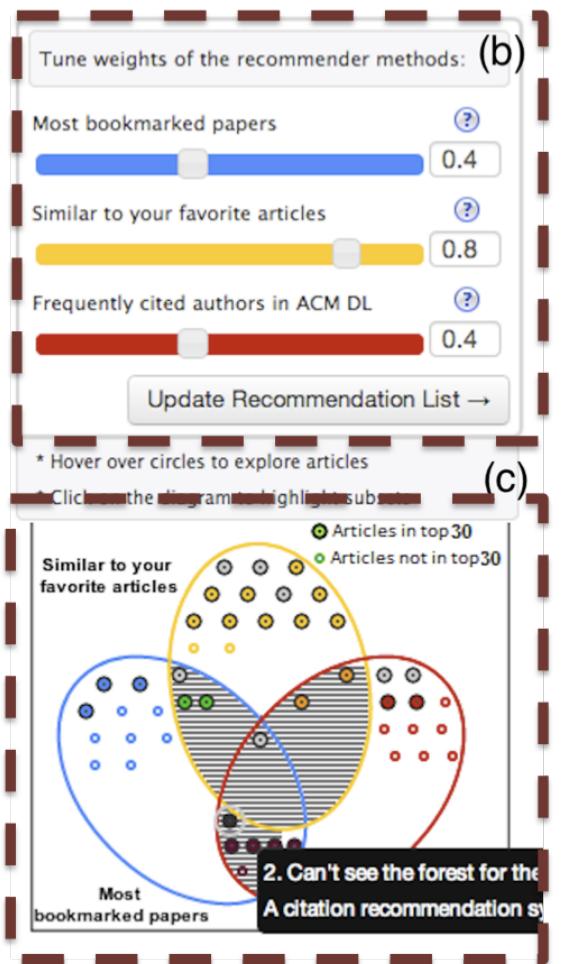
H. Paralela: Mixed III

Traditional Ranked List

Papers sorted by Relevance.
It combines 3 recommendation approaches.

(a)	
2. Can't see the forest for the trees? A citation recommendation system	 [see abstract]
by C. Lee Giles, Cornelia Caragea, Adrian Silvescu, Prasenjit Mitra	
3. When thumbnails are and are not enough: Factors behind users	 [see abstract]
by Dan Albertson	
7. Gendered Artifacts and User Agency	 [see abstract]
by Andrea R. Marshall, Jennifer A. Rode	
8. Two Paths to Motivation through Game Design Elements: Reward-Based Gamification and Meaningful Gamification	 [see abstract]
by Scott Nicholson	
9. Automatic Identifying Search Tactic in Individual Information Seeking: A Hidden Markov Model Approach	 [see abstract]
by Zhen Yue, Shuguang Han, Daqing He	
11. Old Maps and Open Data Networks	 [see abstract]
by Werner Robitz, Carl Lagoze, Bernhard Haslhofer, Keith Newman, Amanda Stefanik	
14. Effects of User Identity Information On Key Answer Outcomes in Social Q&A	 [see abstract]
by Erik Choi, Craig Scott, Chirag Shah	
15. Ebooks and cross generational perceived privacy issues	 [see abstract]
by Jennifer Sue Thiele, Renee Kapusniak	
16. Toward a mesoscopic analysis of the temporal evolution of scientific collaboration networks	

H. Paralela: Mixed IV



Sliders

Allow the user to control the importance of each data source or recommendation method

Interactive Venn Diagram

Allows the user to inspect and to filter papers recommended. Actions available:

- Filter item list by clicking on an area
- Highlight a paper by mouse-over on a circle
- Scroll to paper by clicking on a circle
- Indicate bookmarked papers

H. Paralela: Switching I

- De un grupo de recomendadores, activar un recomendador a la vez.
- Podría ser especialmente útil considerando los learning rate de algunos métodos.
- Ejemplo: Elegir entre un clasificador Bayesiano y un recomendador Item-based como en: Ghazanfar, M., & Prugel-Bennett, A. (2010). An Improved Switching Hybrid Recommender System Using Naive Bayes Classifier and Collaborative Filtering.

$$P(C_j|d) = \frac{P(C_j) \prod_{i=1}^h P(F_i|C_j)}{P(F_1, \dots, F_h)}.$$

H. Paralela: Switching II

Table 1: A comparison of proposed algorithm with existing in terms of cost (based on [31]), accuracy metrics, and coverage

Algorithm	On-line Cost	Best MAE		ROC-Sensitivity		Coverage	
		(ML)	(FT)	(ML)	(FT)	(ML)	(FT)
$UBCF_{DV}$	$O(M^2N) + O(NM)$	0.766	1.441	0.706	0.563	99.424	93.611
IBCF	$O(N^2)$	0.763	1.421	0.733	0.605	99.221	92.312
IDemo4	$O(N^2)$	0.749	1.407	0.739	0.621	99.541	94.435
Rec_{NBCF}	$O(N^2) + O(Mf)$	0.696	1.341	0.778	0.657	100	99.992
NB	$O(Mf)$	0.808	1.462	0.703	0.571	100	99.992
NH	$O(N^2) + O(Mf)$	0.785	1.438	0.712	0.586	100	99.992
CB	$O(M^2N) + O(NM) + O(Mf)$	0.721	1.378	0.741	0.611	100	99.995

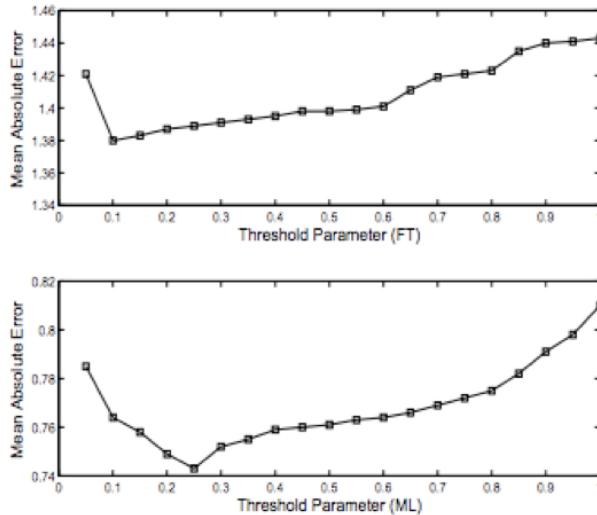
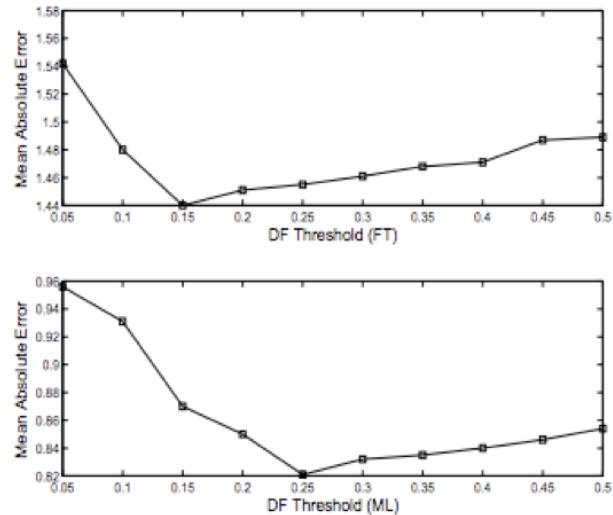


Figure 1: Determining the optimal value of DF .

Figure 2: Determining the optimal value of α .

Hibridización Pipeline

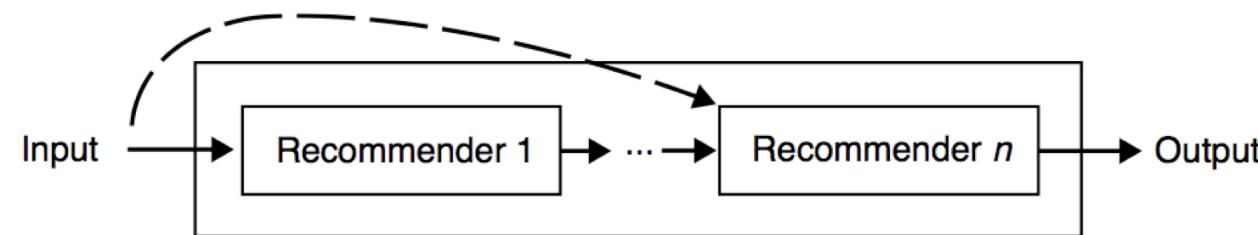


Figure 5.4. Pipelined hybridization design.

Dos mecanismos principales:

- Cascade
- Meta-Level

H. Pipeline: Cascade

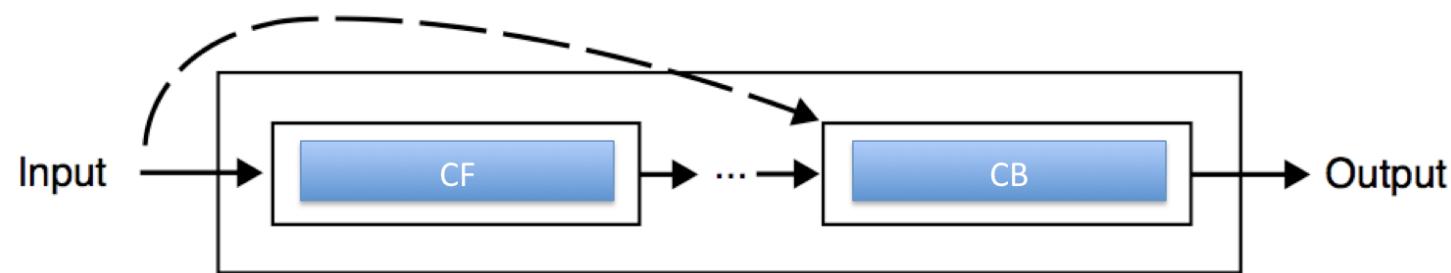
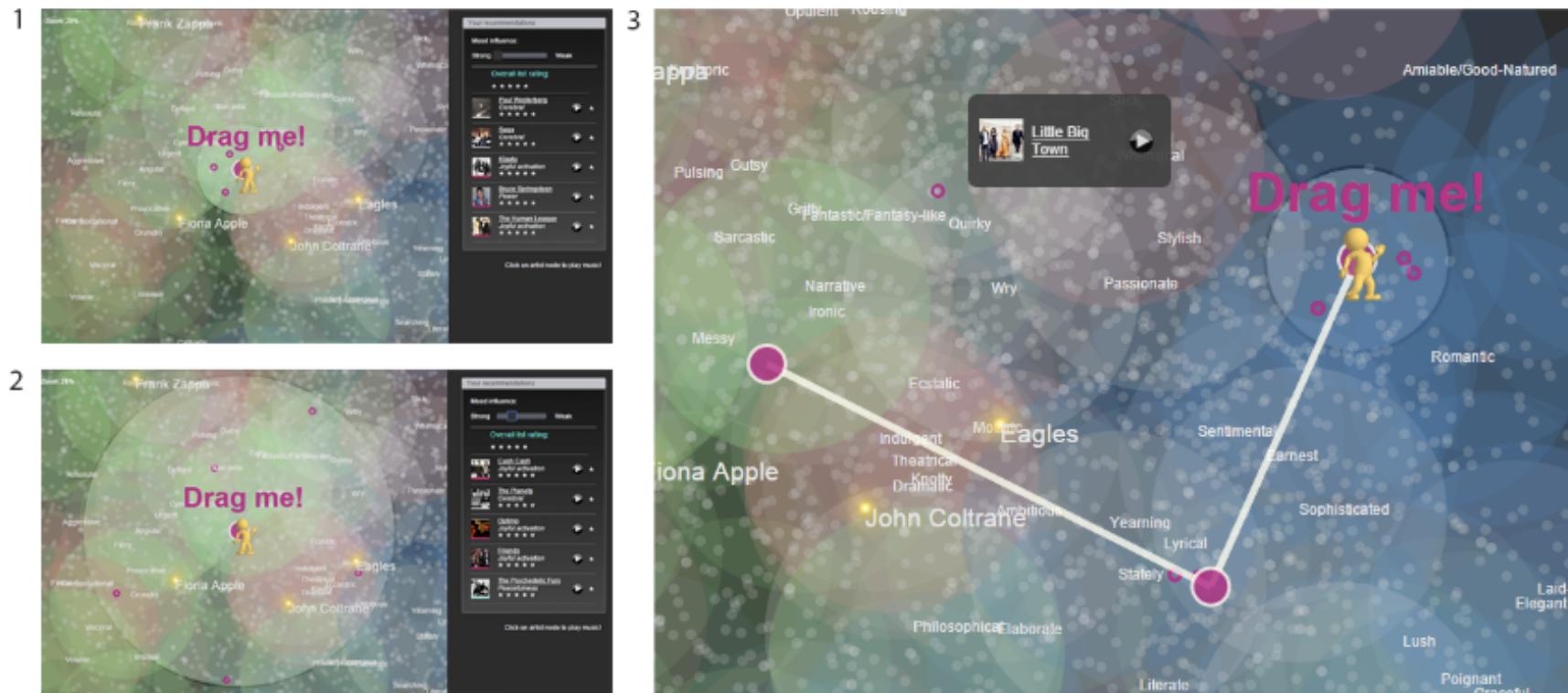


Figure 5.4. Pipelined hybridization design.

H. Pipeline: Cascade (Moodplay)



H. Pipeline: Cascade (Moodplay)

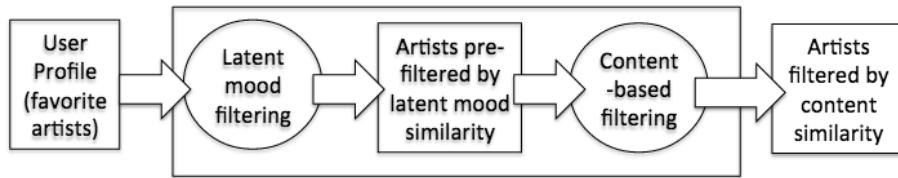
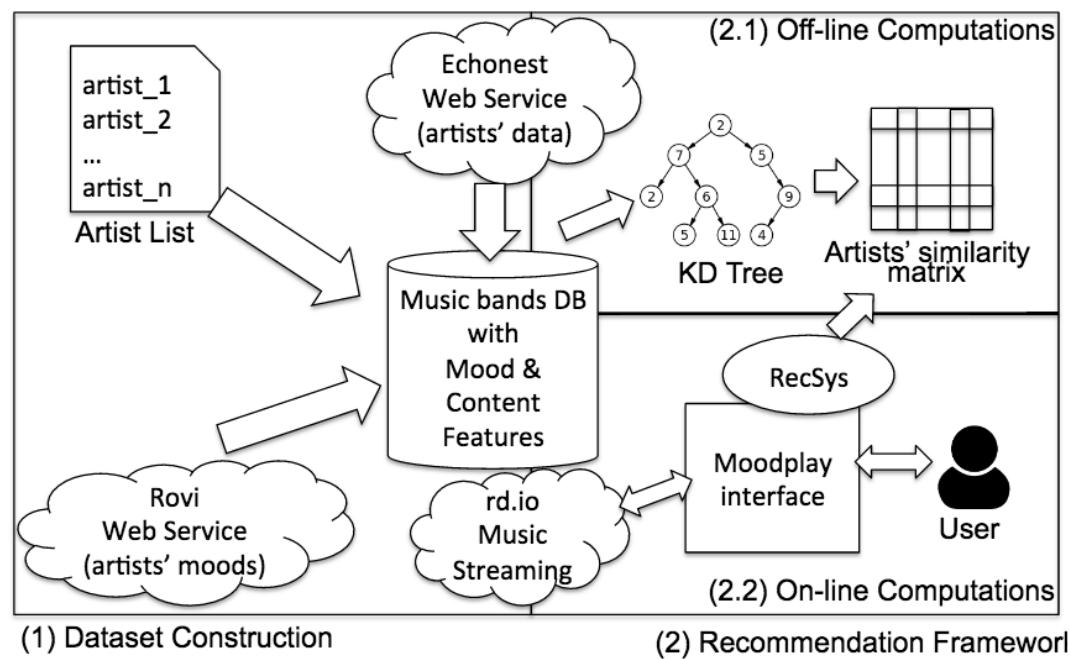
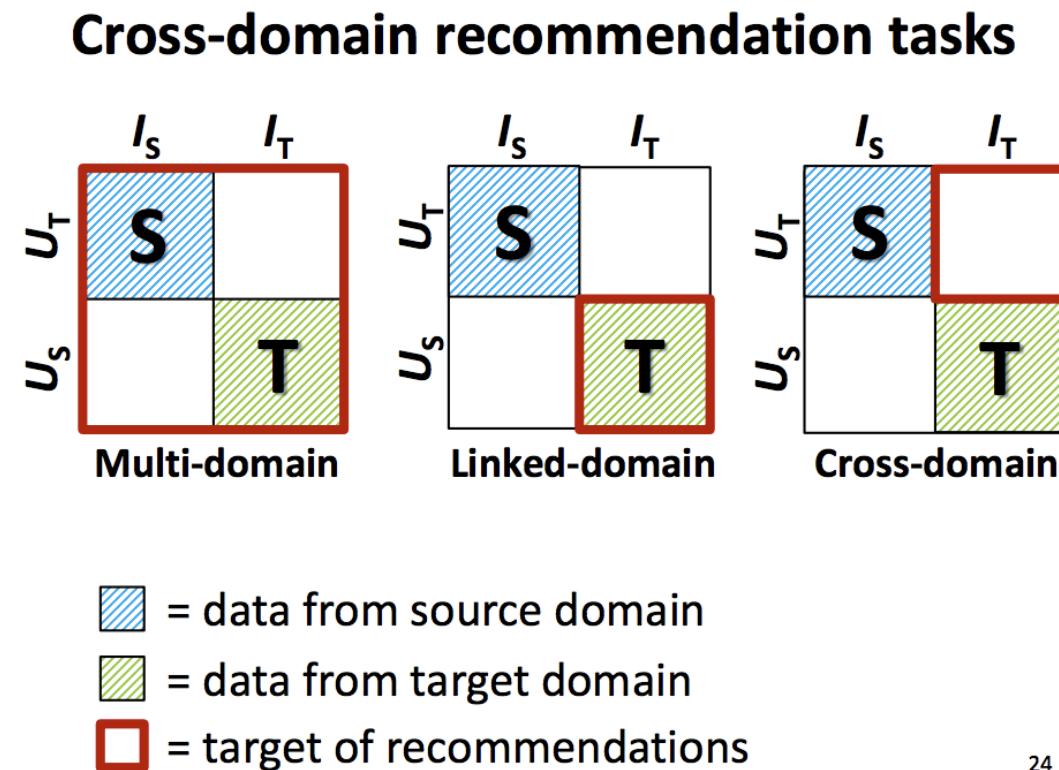


Fig. 4 Schematic representing our hybrid cascading recommender which pre-filters based on mood larity and then post-filters based on content similarity.



H. Pipeline: Meta-Level

El modelo aprendido por un recomendador es usado para un segundo recomendador.
Transfer Learning:



24

Tutorial on Cross-domain recommender systems http://recsys.acm.org/wp-content/uploads/2014/10/recsys2014_tutorial_cross_domain.pdf

Blending: Learn from the masters

- Netflix prize winners (BigChaos)
 - Jahrer, M., Töscher, A., & Legenstein, R. (2010, July). Combining predictions for accurate recommender systems. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 693-702). ACM.
 - Töscher, A., Jahrer, M., & Bell, R. M. (2009). The bigchaos solution to the netflix grand prize. Netflix prize documentation, 1-52.
 - Töscher, A., & Jahrer, M. (2008). The bigchaos solution to the netflix prize 2008. Netflix Prize, Report.

Blending: Learn from the masters

- Similar problem: CTR
 - Jahrer, M., Toscher, A., Lee, J. Y., Deng, J., Zhang, H., & Spoelstra, J. (2012). Ensemble of collaborative filtering and feature engineered models for click through rate prediction. In KDDCup Workshop.

Referencias

- Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4), 331-370.
- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6), 734-749.
- Burke, R. (2007). Hybrid web recommender systems. In *The adaptive web* (pp. 377-408). Springer Berlin Heidelberg.
- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). *Recommender systems: an introduction*. Cambridge University Press. Chicago

