Hybrid Recommendation

Peter Brusilovsky
with slides of Danielle Lee

IS2480 Adaptive Information Systems

Where we are?

	Search	Navigation	Recommendation
Content-based			
Semantics / Metadata			
Social			

Three basic recommendation engines

- **Collaborative Filtering**: exploiting other likely-minded community data to derive recommendations
  - Effective, Novel and Serendipitous recommendations
  - Data Sparsity, cold-start problem and ad-hoc users
- **Content-based approach**: relying on product (information) features and textual descriptions
- **Knowledge-based approach** : reasoning on explicit knowledge models from the domain
  - Ability to generate recommendation with a small set of user preference and suggest reasonable recommendations
  - Easy to generate too obvious or boring recommendation and plasticity problems.
- Each engine also have variations
  - Content vs. metadata in CBF
  - Peers vs. friends in CF

Input Data Requirements of Recommendation Techniques

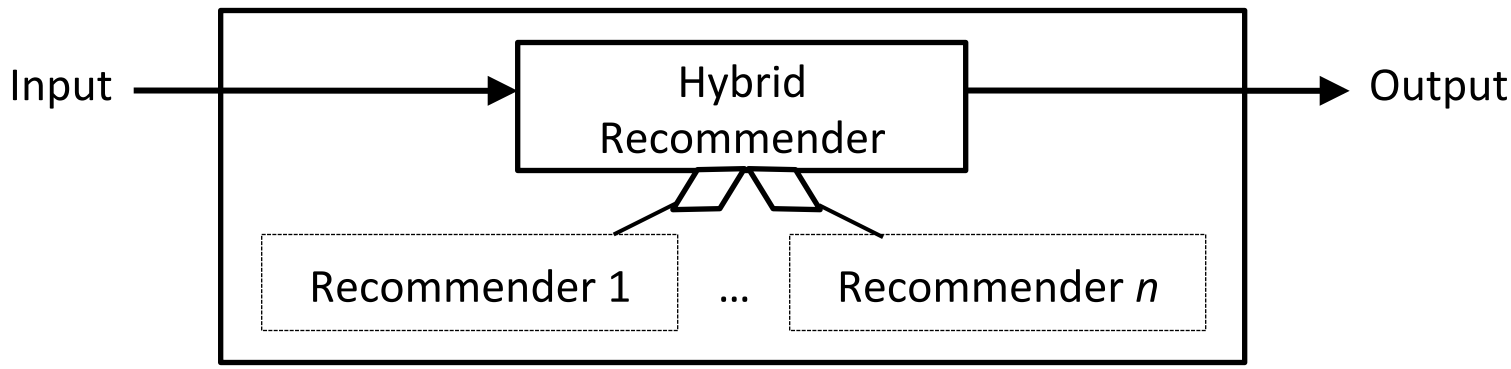
	User Profile & Contextual Parameters	Community Data	Product Features	Knowledge models
Collaborative Filtering	Yes	Yes	No	No
Content-based	Yes	No	Yes	No
Knowledge-based	Yes	No	Yes	Yes

Different engines and their variations typically use difference sources of data. It could be wise to combine the approaches to use more data

Hybridization Designs

- Monolithic Hybridization
  - Incorporating aspects of several recommendation strategies in one algorithm implementation
- Parallelized Hybridization
  - Operating independently of one another and produce separate recommendation lists. Then their output is combined into a final set of recommendations
- Pipelined Hybridization
  - Several recommender systems are joined together in a pipeline architecture. The output of one recommender becomes part of the input of the subsequent one.

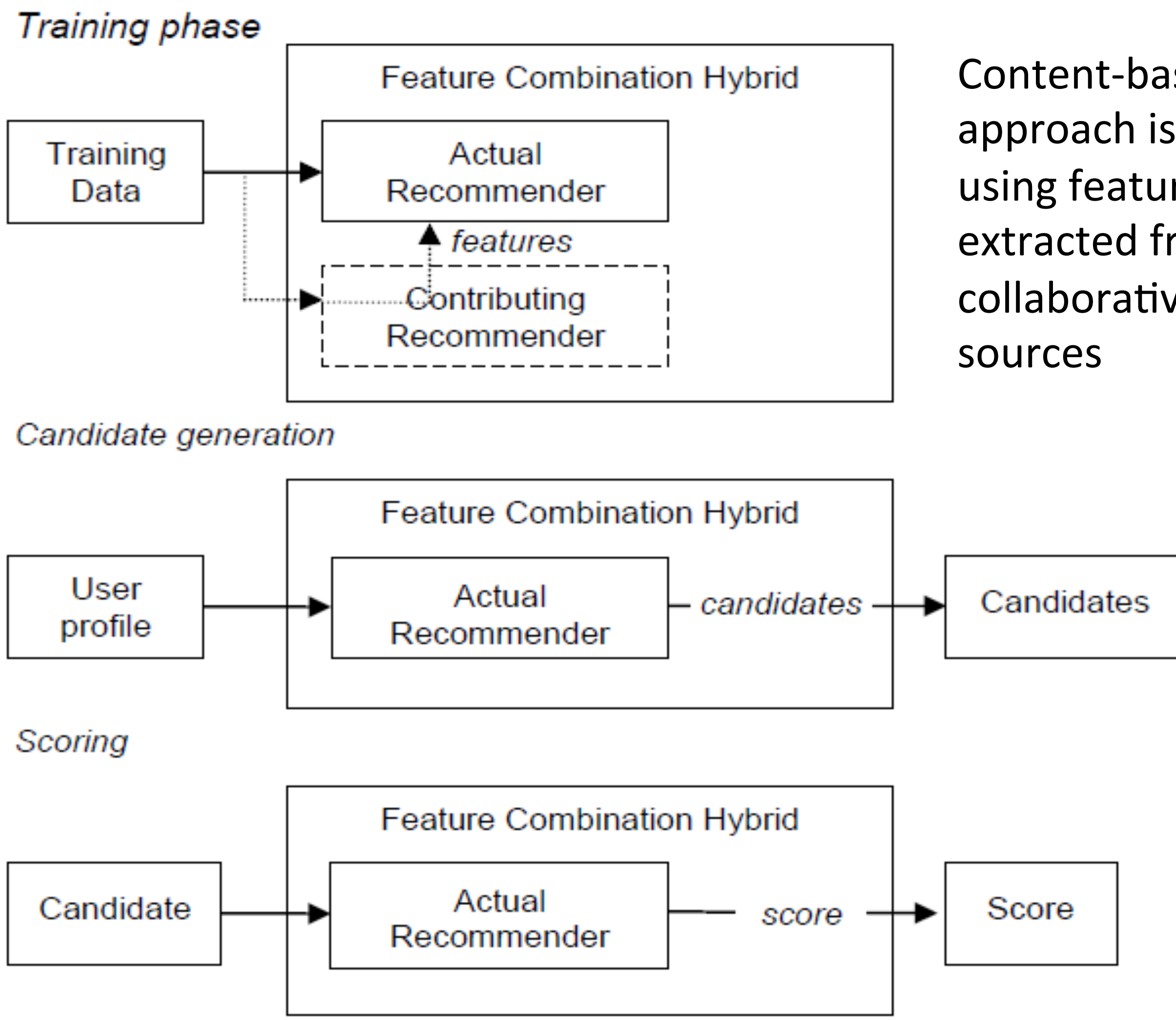
Monolithic Hybridization



- Built-in modification of recommendation algorithm to exploit different types of input data
- Apply one approach (i.e. CBF) but enhance with the knowledge sources that are typically used by other (CoF)
- Feature combination hybrids
  - Ex) Basu, et al. (1998), Zanker and Jessenitschnig (2009), Pazzani (1999)
- Feature augmentation hybrids
  - Melville, et al. (2002), Mooney and Roy (1999), and Torres et al. (2004)

Monolithic Hybridization

- Feature combination hybrids



Content-based approach is trained using features extracted from collaborative sources

Example (1)

User	Item1	Item2	Item3	Item4	Item5
Alice		1		1	
User1		1	1		1
User2	1	1			1
User3	1		1		
User4					1

Item	Genre
Item1	Romance
Item2	Mystery
Item3	Mystery
Item4	Mystery
Item5	Fiction

Example (1)

Feature	Alice	User1	User2	User3	User4
User likes many *mystery* books	true	true			
User likes some *mystery* books			true	true	
User likes many *romance* books					
User likes some *romance* books			true	true	
User likes many *fiction* books					
User likes some *fiction* books		true	true		true

Legend: If a user bought mainly books of genre *X* (two-thirds of the total purchases and at least two books), we say that 'Users likes many X books'

Example (2)

	R nav	R view	R ctx	R buy
Alice	n3, n4	i5	k5	null
User1	n1, n5	i3, i5	k5	i1
User2	n3, n4	i3, i5, i7	null	i3
User3	n2, n3, n4	i2, i4, i5	2, k4	i4

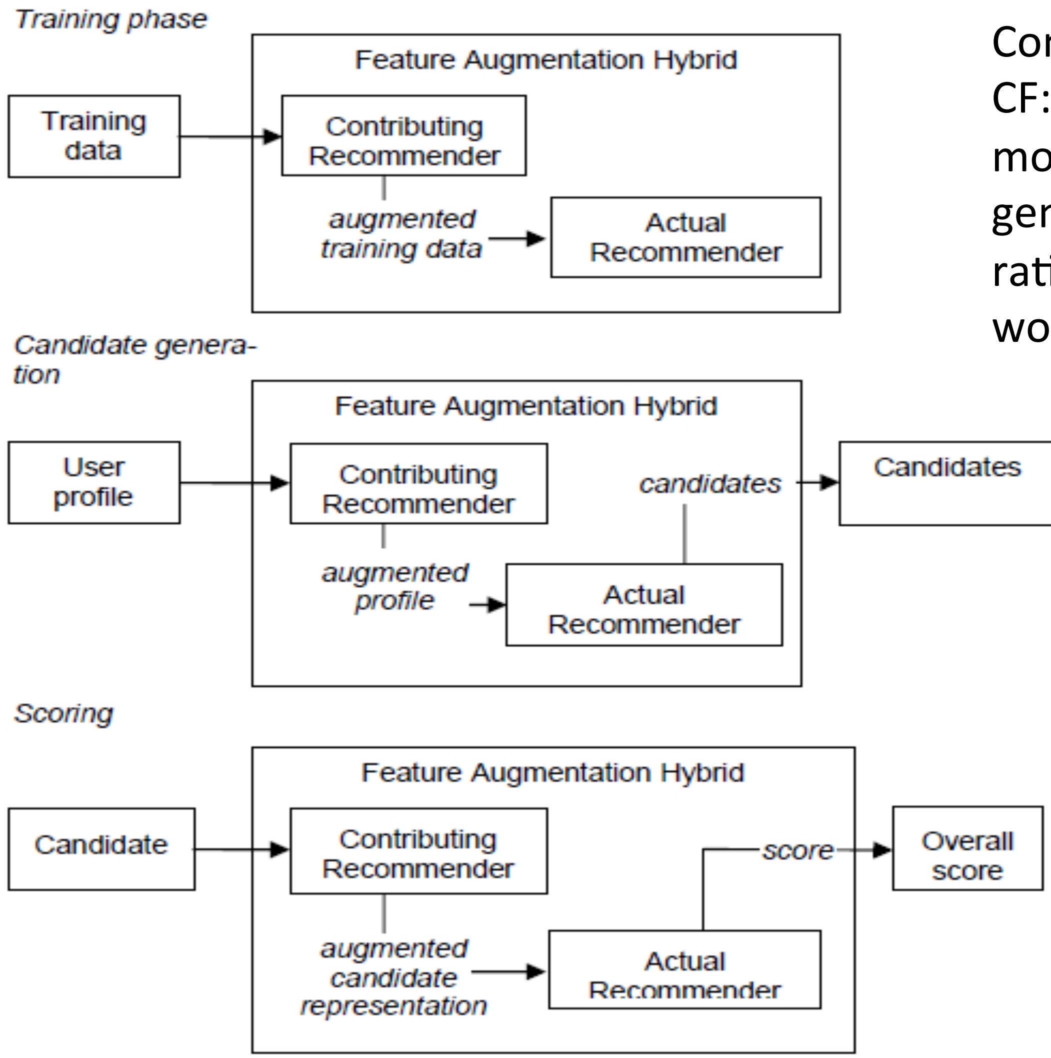
Precedence rules: (R buy, R ctx) - R view - R nav

Example (3)

- Elicitation of user feedback and collaborative filtering
  - *Price should be less than the price for item a.*

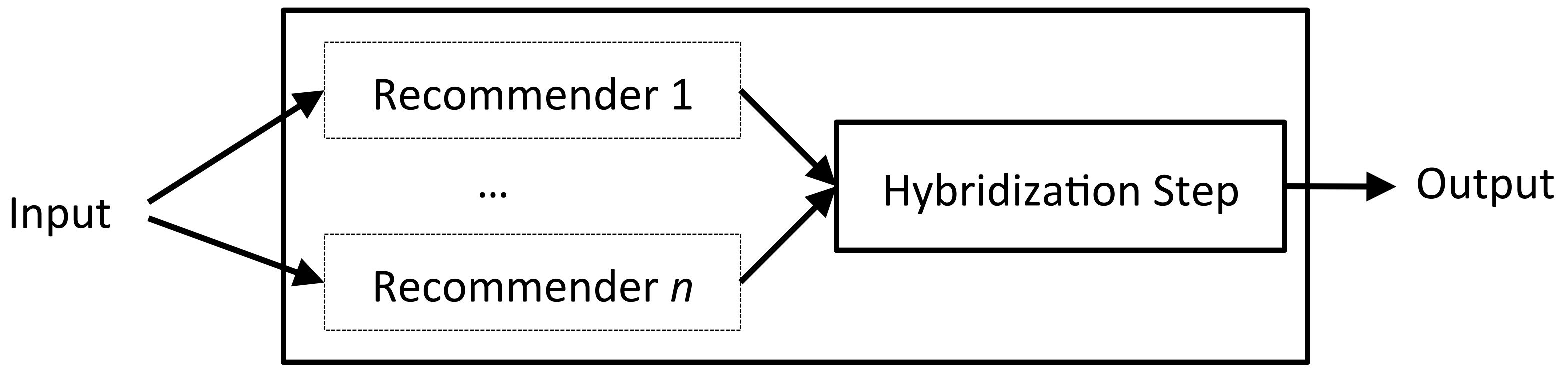
Monolithic Hybridization

- Feature augmentation hybrids



Content-boosted CF: content-based model used to generate missed ratings. Then CF works

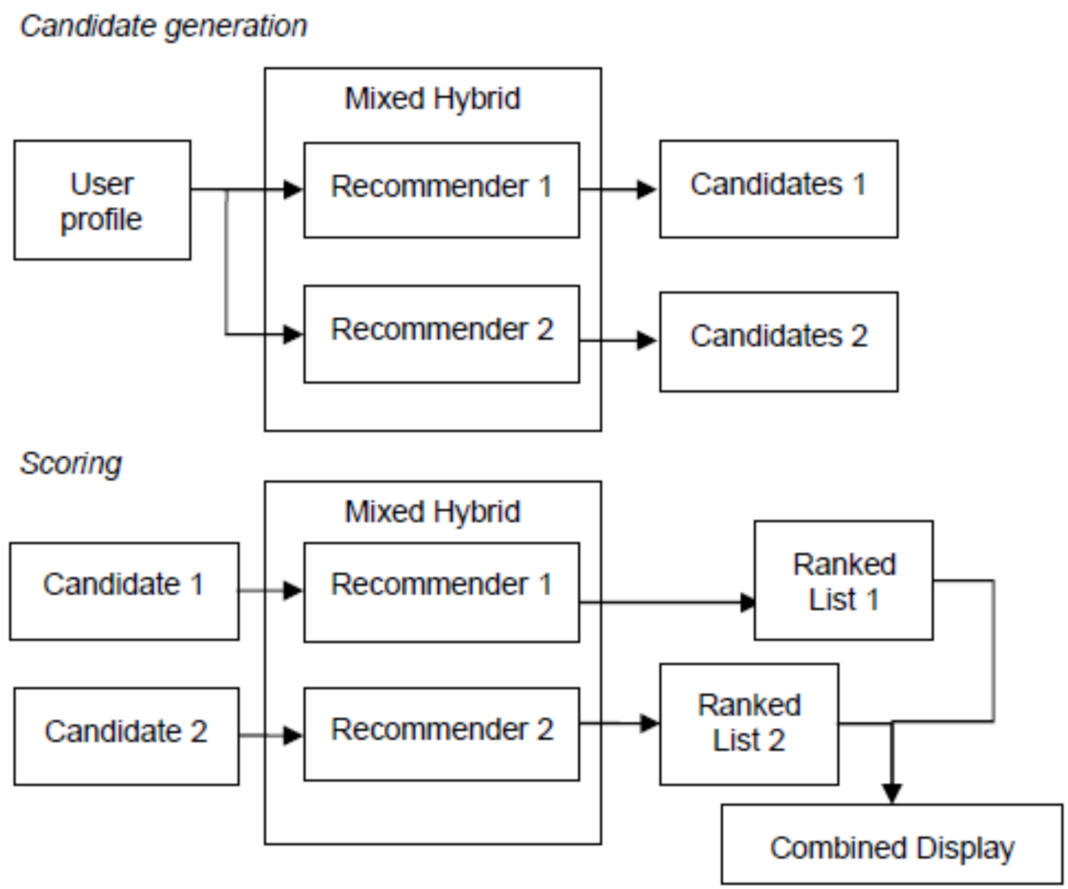
Parallelized Hybridization



- Employ several recommenders side by side and employ a specific hybridization technique to aggregate the outputs.
- Mixed Hybrids
  - Cotter & Smyth (2000), Zanker, et al. (2007)
- Weighted Hybrids
  - Zanker and Jessenitschnig (2009), Claypool, et al. (1999)
- Switching Hybrids
  - Zanker and Jessenitschnig (2009), van Setten (2005)

Parallelized Hybridization

- Mixed Hybrid: combines results of different recommenders at user interface level



Example of Combination

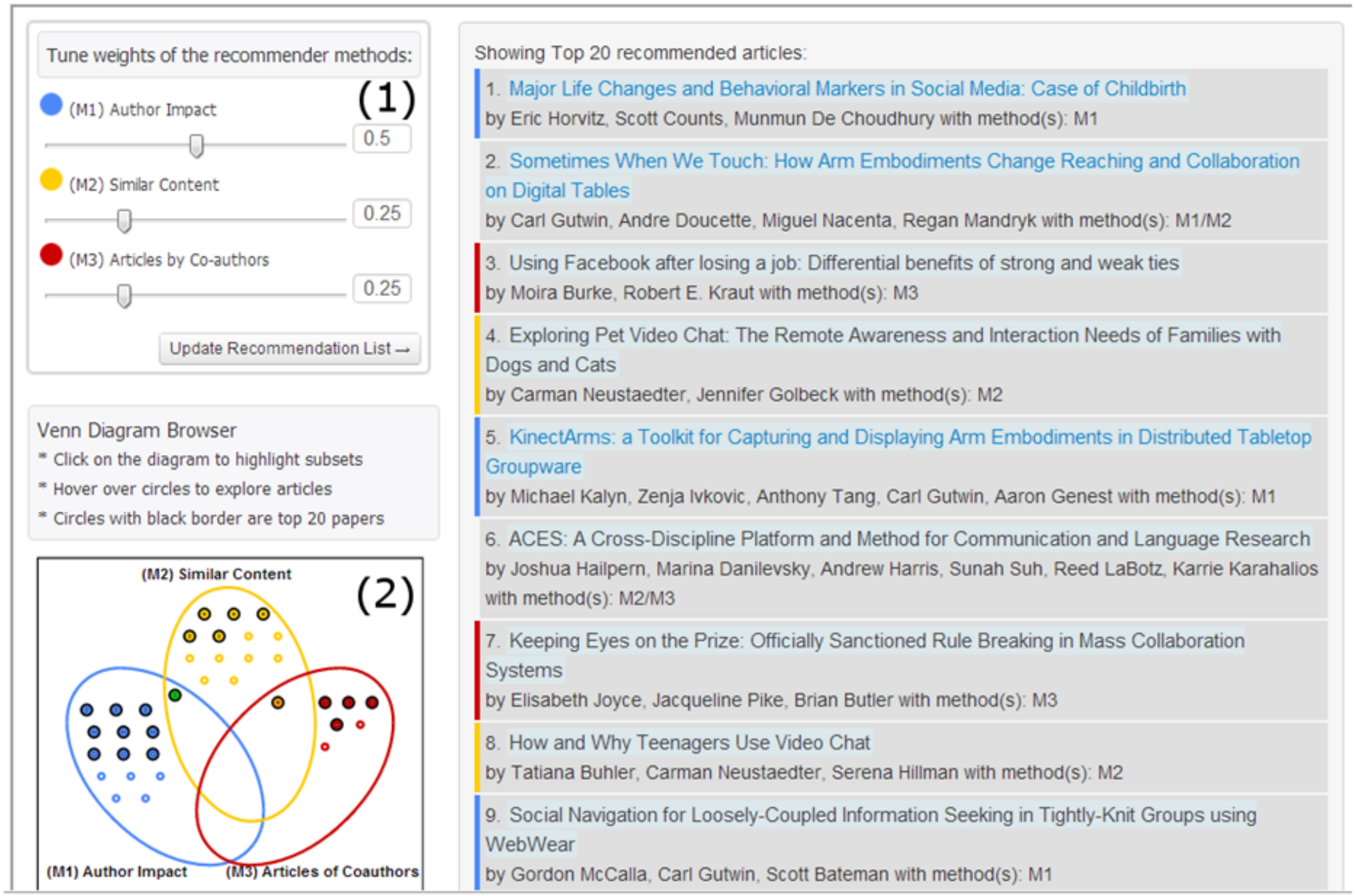
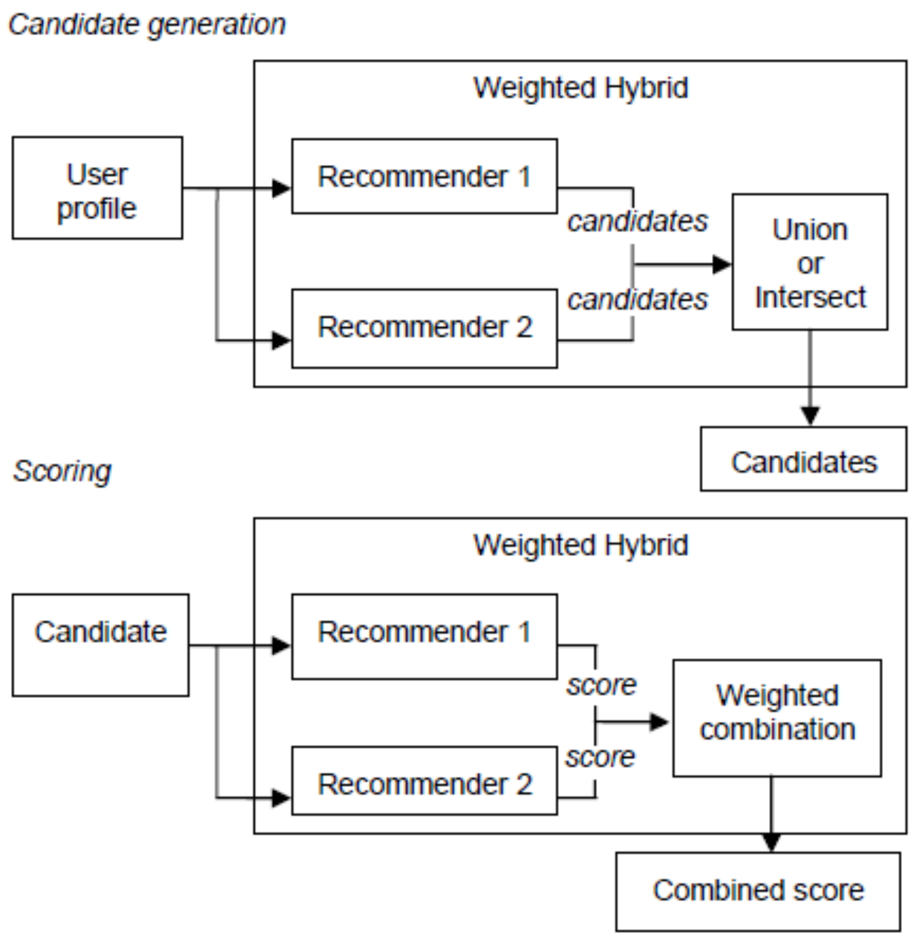


Image showing the condition of an interactive controllable interface. In addition to browsing a list the articles, the user can control (sliders at the top

Parallelized Hybridization

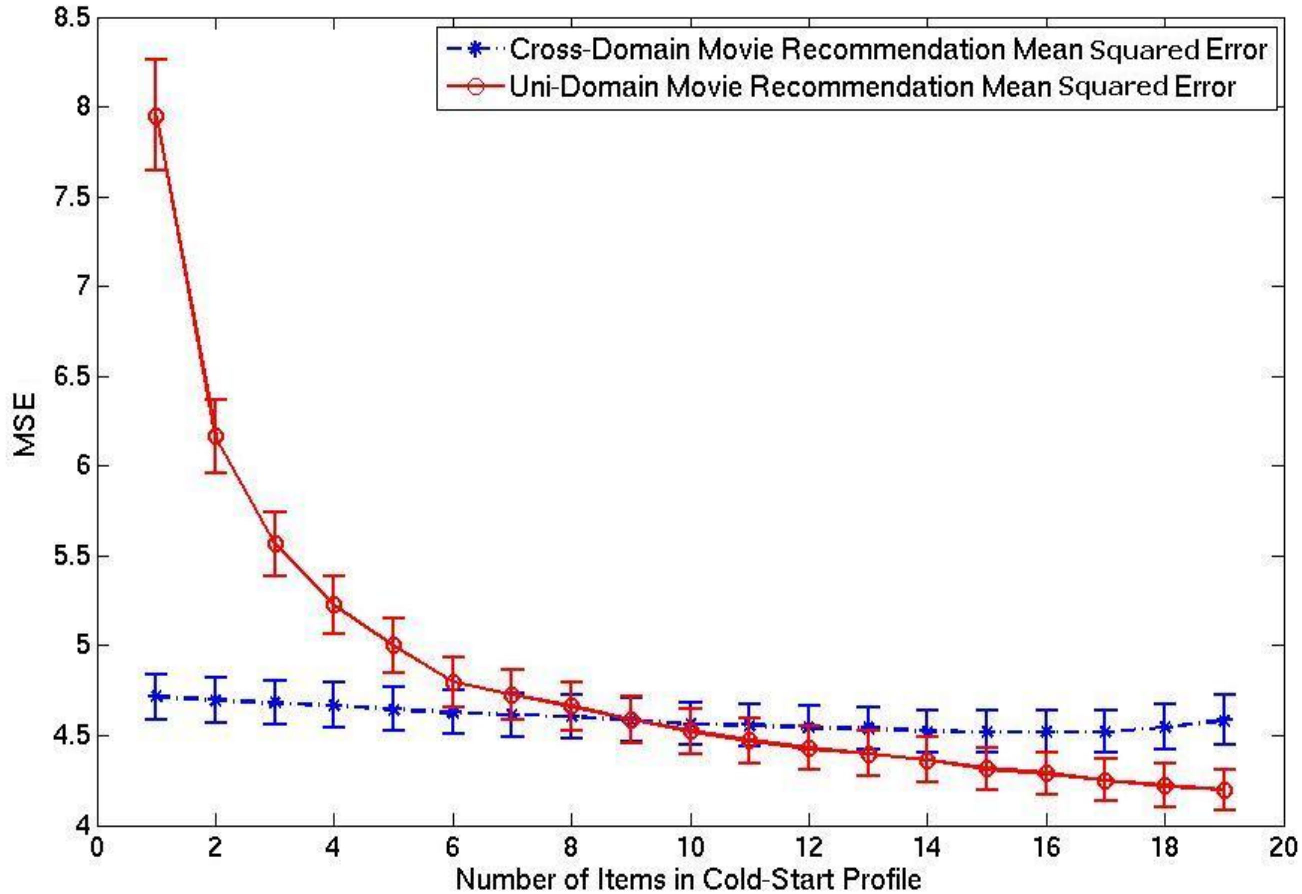
- Weighted Hybrids: Combines recommendations by computing weighted sums of their scores



Parallelized Hybridization

	rec1 score	rec1 rank	rec2 score	rec2 rank	recw score	recw 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	

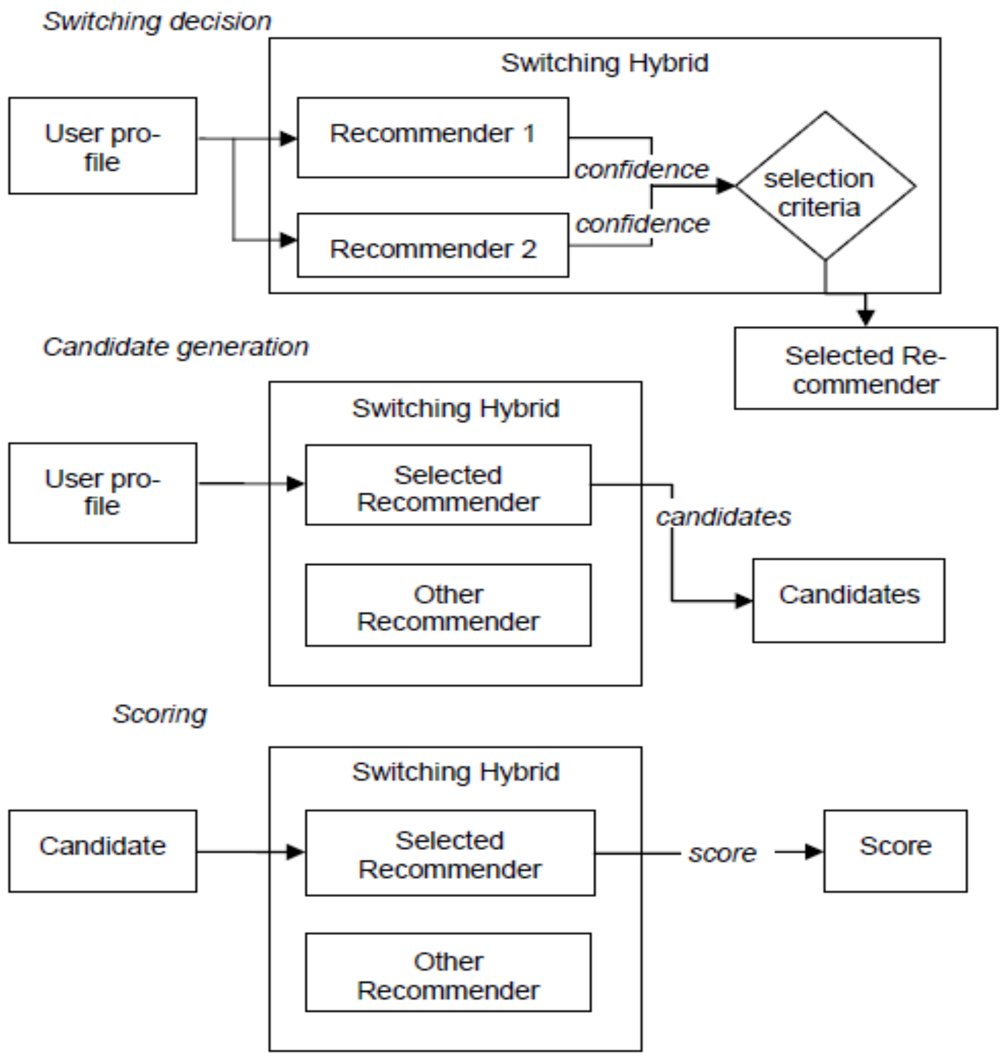
Parallelized Hybridization



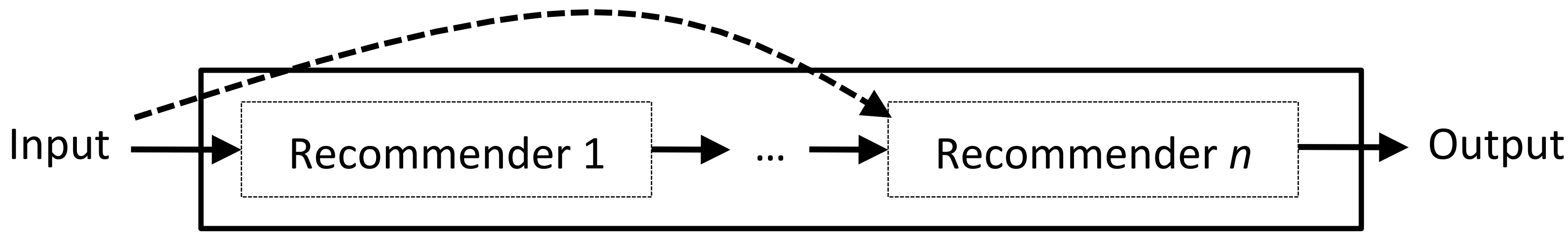
Why switching might be better than weighting?

Parallelized Hybridization

- Switching hybrids



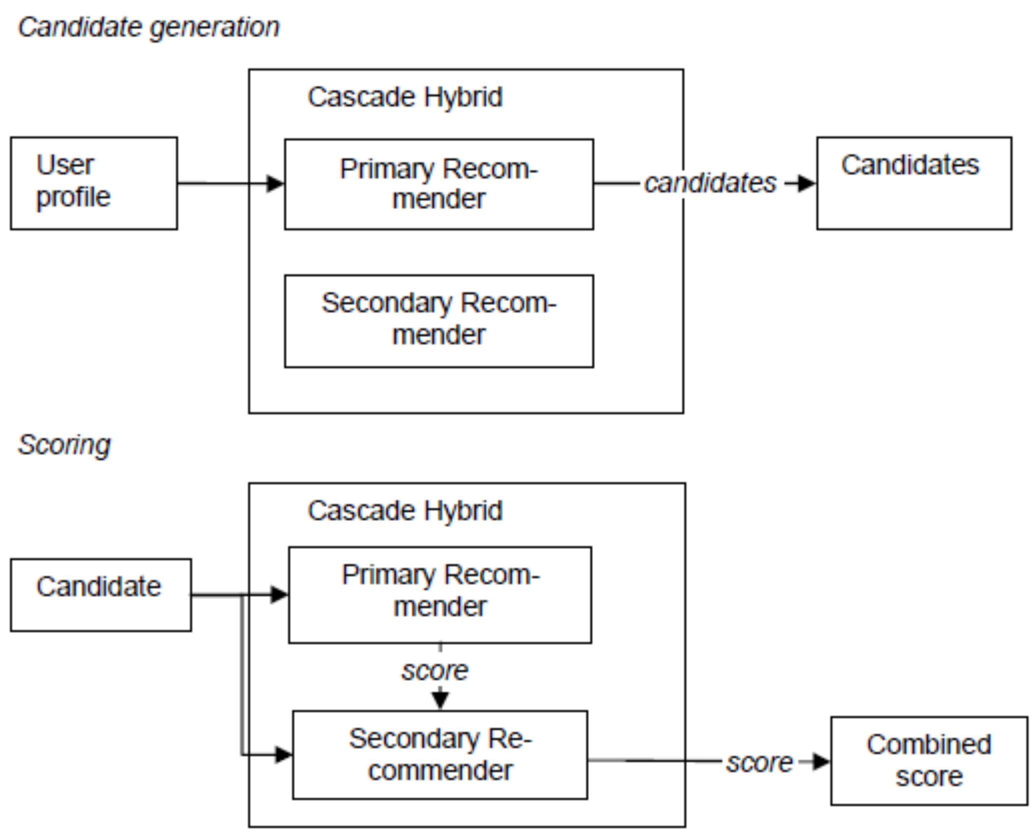
Pipelined Hybridization



- A staged process in which several techniques sequentially build on each other before the final one produces recommendations
- Cascade Hybrids
  – Zanker and Jessenitschnig (2009)
- Meta-level Hybrids
  – Zanker (2008), Pazzani (1999)

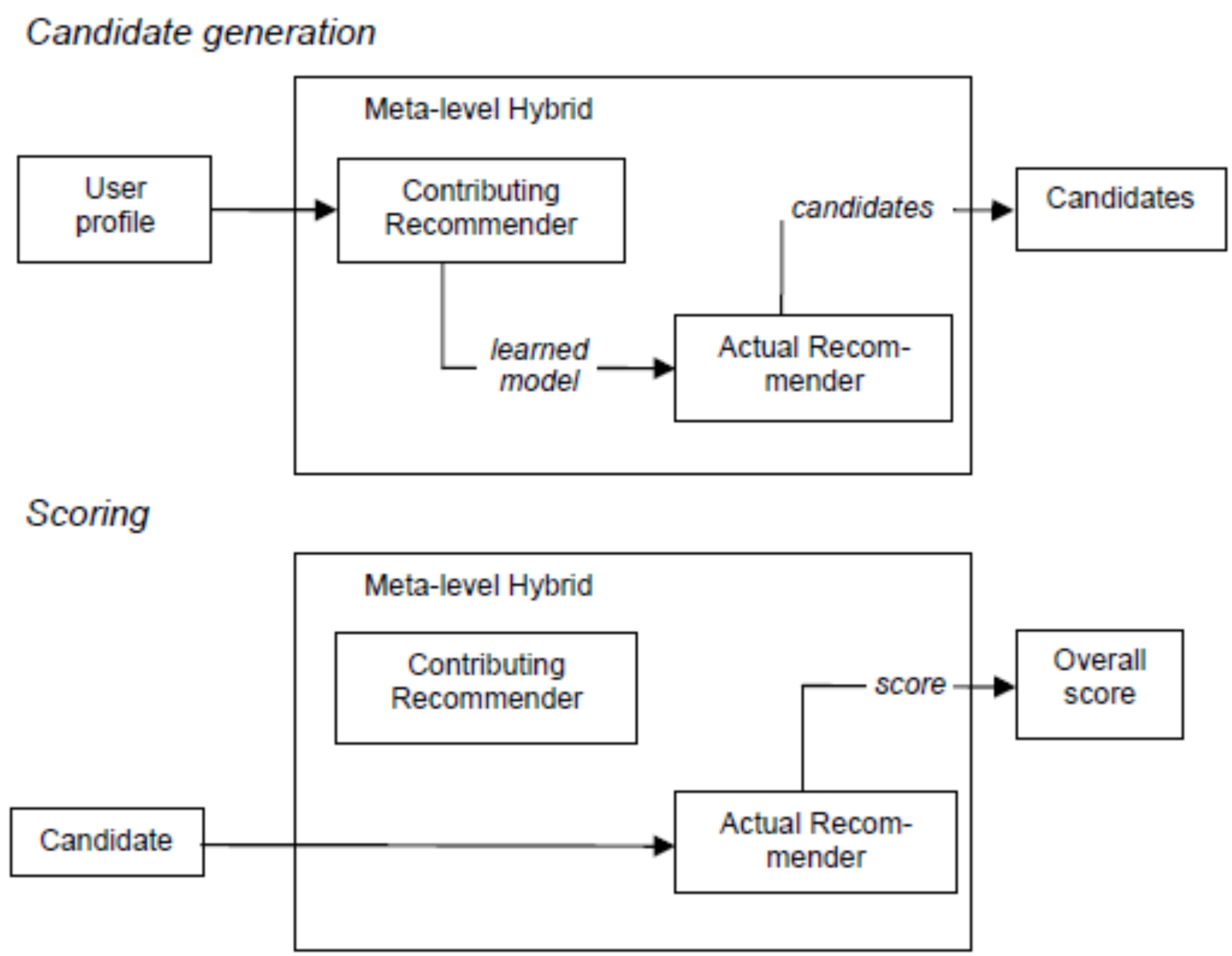
Pipelined Hybridization

- Cascade hybrids: based on a sequenced order of techniques.



Pipelined Hybridization

- Meta-Level Hybrids: one recommender builds a model that is exploited by the principal recommender



Hybridization Summary

	Weight.	Mixed	Switch.	FC	Cascade	FA	Meta
CF/CN							
CF/DM							
CF/KB							
CN/CF							
CN/DM							
CN/KB							
DM/CF							
DM/CN							
DM/KB							
KB/CF							
KB/CN							
KB/DM							

FC = Feature Combination, FA = Feature Augmentation
CF = collaborative, CN = content-based, DM = demographic, KB = knowledge-based

	Redundant
	Not possible
	Existing implementation