Probabilistic relational models for recommendation systems: review and comparative study

Mouna Ben Ishak,Rajani Chulyadyo, Philippe Leray,
Ahmed Abdelwahab, Miriam Ramirez and Nahla Ben Amor

UNIVERSITÉ DE NANTES
POLYTECH
lina
DUKe

UNIVERSITÉ DE TUNIS
LARODEC

Motivation

Recommender Systems

- Recommender Systems (RSs) are software tools and techniques providing suggestions for items to be of use to a particular user within a specific domain.
- *Examples :* Amazon, YouTube, etc.

Motivation

Recommender Systems

- Recommender Systems (RSs) are software tools and techniques providing suggestions for items to be of use to a particular user within a specific domain.
- *Examples :* Amazon, YouTube, etc.

Recommendation techniques

- Collaborative filtering
- Content-based
- Demographic
- ...

Motivation

Major Challenges

- Scalability - undegraded performance with growing data
- Data sparsity - Sparse User-Item matrix
- Cold start problem - ability to perform recommendation for new user/item or in a new system

Existing techniques drawbacks

Motivation

Major Challenges

Existing techniques drawbacks

- rely on propositional data representation.
- do not exploit all the available information :
  - Collaborative-based approaches are based on the user-item matrix.
  - Content-based approaches are based on the users profiles.
  - ...

Hybrid methods

Motivation

Major Challenges

Existing techniques drawbacks

Hybrid methods

Showed better performance than standard techniques.

Motivation

Motivation

- Data mining approaches are widely used to perform recommendation, e.g., Bayesian networks.

Motivation

- Data mining approaches are widely used to perform recommendation, e.g., Bayesian networks.
- Real world domains are often designed using relational representation allowing interaction of heterogeneous entities.

Motivation

- Data mining approaches are widely used to perform recommendation, e.g., Bayesian networks.
- Real world domains are often designed using relational representation allowing interaction of heterogeneous entities.
- More expressive classes of directed graphical models have emerged to extract statistical patterns from relational data.

Motivation

- Data mining approaches are widely used to perform recommendation, e.g., Bayesian networks.
- Real world domains are often designed using relational representation allowing interaction of heterogeneous entities.
- More expressive classes of directed graphical models have emerged to extract statistical patterns from relational data.

Relational datamining models

Performing recommendation using a relational model :
Probabilistic Relational Models (Koller and Pfeffer 1998)

Outline

Outline...

1. Probabilistic relational models
2. Recommendation using PRMs
3. Implementation
4. Conclusion & perspective

Bayesian networks *(Pearl 1988)*

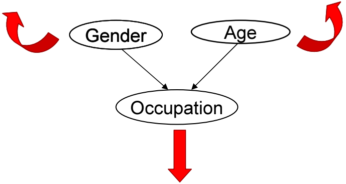
Definition

A BN is composed of 2 components :

- Qualitative : **DAG** over a set of variables.
- Quantitative : set of *CPTs* = *Conditional Probability Table* of each variable given its parents in the graph.

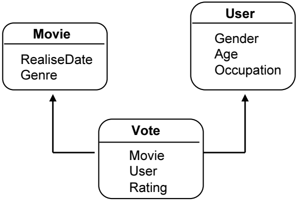
Gender	
M	F
0.4	0.6

Age		
Low	Middle	High
0.4	0.3	0.3



Age / Gender	Occupation		
	Oc1	Oc2	Oc3
Low,F	0.5	0.1	0.4
Low,M	0.3	0.5	0.2
Middle,F	0.2	0.4	0.4
Middle,M	0.9	0.1	0
High,F	0.3	0.5	0.2
High,M	0.2	0.3	0.5

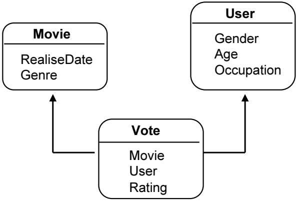
Probabilistic Relational Models *(Koller and Pfeffer 1998)*



An example of a Relational Schema

A relational schema $\mathcal{R}$

Probabilistic Relational Models *(Koller and Pfeffer 1998)*

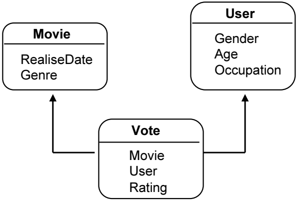


An example of a Relational Schema

A relational schema $\mathcal{R}$

- Classes + relational variables.

Probabilistic Relational Models *(Koller and Pfeffer 1998)*

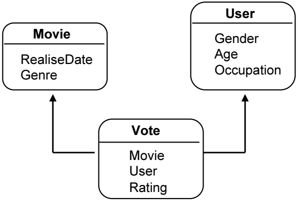


An example of a Relational Schema

A relational schema $\mathcal{R}$

- Classes + relational variables.
- Reference slots (e.g., *Vote.Movie*, *Vote.User*).

Probabilistic Relational Models *(Koller and Pfeffer 1998)*



An example of a Relational Schema

A relational schema $\mathcal{R}$

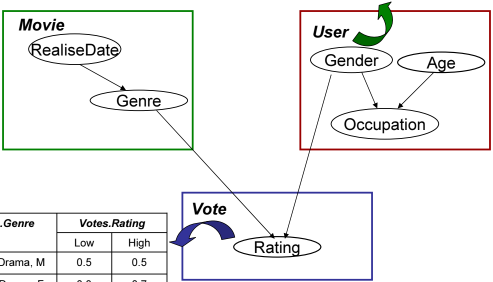
- Classes + relational variables.
- Reference slots (e.g., $Vote.Movie$, $Vote.User$).
- Slot chain = A sequence of reference slots
  - Allow to walk in the relational schema to create new variables (e.g., $Vote.User.User^{-1}.Movie$: all the movies voted by a particular user.)

Probabilistic Relational Models *(Koller and Pfeffer 1998)*

Definition

Aggregators

User.Gender	
M	F
0.4	0.6



Movie.Genre / *User.Gender*	*Votes.Rating* Low	High
Drama, M	0.5	0.5
Drama, F	0.3	0.7
Horror, M	0.2	0.8
Horror, F	0.9	0.1
Comedy, M	0.5	0.5
Comedy, F	0.6	0.4

Probabilistic Relational Models *(Koller and Pfeffer 1998)*

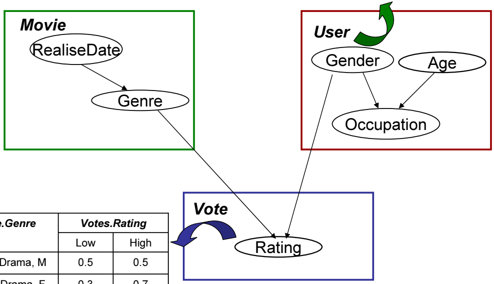
Definition

A PRM Π for a relational schema R :

- A qualitative dependency structure $\mathcal{S}$ (with possible long **slot chains** and **aggregation functions**).
- A set of parameters $\theta_{\mathcal{S}}$.

Aggregators

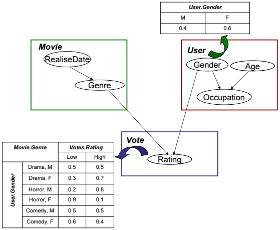
User.Gender	
M	F
0.4	0.6



Movie.Genre		***Votes.Rating***	
		Low	High
User.Gender	Drama, M	0.5	0.5
	Drama, F	0.3	0.7
	Horror, M	0.2	0.8
	Horror, F	0.9	0.1
	Comedy, M	0.5	0.5
	Comedy, F	0.6	0.4

Probabilistic Relational Models *(Koller and Pfeffer 1998)*

Definition



Movie.Genre, User.Gender	Votes.Rating Low	Votes.Rating High
Drama, M	0.5	0.5
Drama, F	0.3	0.7
Horror, M	0.2	0.8
Horror, F	0.9	0.1
Comedy, M	0.5	0.5
Comedy, F	0.6	0.4

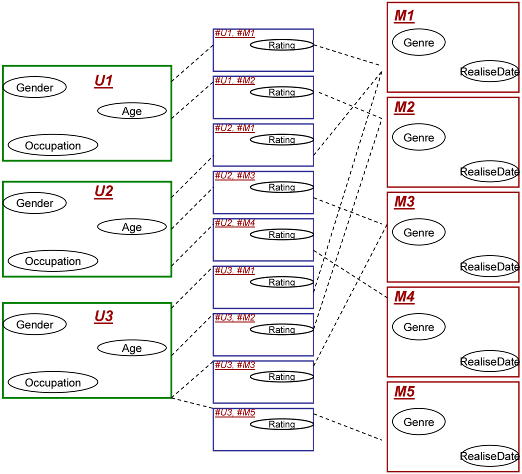
Aggregators

- $\gamma(\mathit{Vote.User.User}^{-1}.\mathit{Movie.genre}) \rightarrow \mathit{Vote.rating}$: All the gender of movies voted by a particular user
  - A finite set, not known in advance to create the conditional probability table.
  - Has to be substitute by an aggregated value (e.g. ($\gamma = \mathit{MODE}$).

Ground Bayesian Network (1/2)

GBN

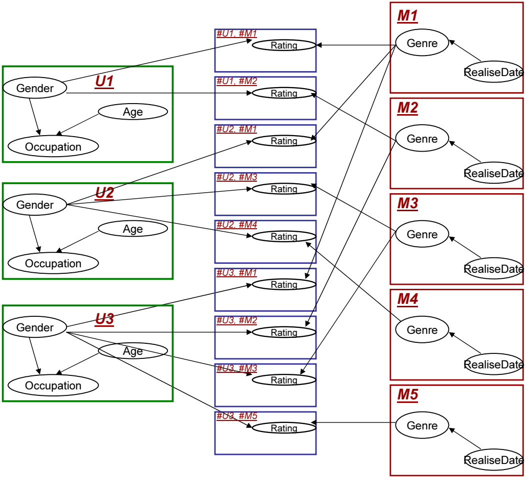
Varies with the relational skeleton that instantiates the model.



Ground Bayesian Network (2/2)

GBN

Varies with the relational skeleton that instantiates the model.



PRMs learning

Close to BN learning.

PRMs learning

Close to BN learning.

- Parameters learning

PRMs learning

Close to BN learning.

- Parameters learning
- Structure learning NP-Hard
  - Probabilistic dependencies may be between two attributes from either the same class or the classes reachable through reference slots.
  - Constraint on slot chain length for adding dependencies.

PRMs learning

Existing methods

Only few works have focused on PRMs structure learning [*RGS (Getoor et al. 2001), RPC, RCD (Maier et al. 2010, 2013)*].

PRMs learning

Existing methods

Only few works have focused on PRMs structure learning [*RGS (Getoor et al. 2001), RPC, RCD (Maier et al. 2010, 2013)*].

Critics

- Lack of evaluation process, in a common framework.
- Absence of relational benchmarks to evaluate the effectiveness of the proposed approaches.
- Non-existence of random generation process for generating synthetic data.

Outline...

PRM for recommendation

- Recommendation = predicting the relationships between users and items.

PRM for recommendation

- Recommendation = predicting the relationships between users and items.
- Once learnt, PRM can be used to predict new relations without going through existing relations. Makes it scalable.

PRM for recommendation

- Recommendation = predicting the relationships between users and items.
- Once learnt, PRM can be used to predict new relations without going through existing relations. Makes it scalable.
- Learns a model directly from relational data. Data sparsity is not an issue.

Related work

Existing approaches

- *Getoor et al., 1999* proposed the idea of parameter learning from incomplete data.
- *Newton et al., 2004* defined hPRMs and used it to perform recommendation.
- *Huang et al., 2005* proposed a unified recommender system based on PRM.
- *Gao et al., 2007* Combine PRMs with other machine learning method.

Critics

Related work

Existing approaches

Critics

- PRMs are either designed for a specific recommendation task or used in a highly basic form due to the intensive computation required by the model estimation process.

Our objective

Related work

Existing approaches

Critics

Our objective

Provide efficient PRMs learning techniques : illustrate our development on the recommendation task.

Steps of our recommendation Process

Adaptation to *Huang et al.'s* Unified Recommendation Framework

- Combines Collaborative, content-based and demographic filtering into a single framework.
- Our focus on rating data.

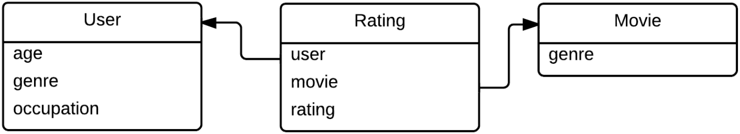


FIGURE : A relational schema for MovieLens dataset

Steps of our recommendation Process

- Relational attributes generation

Steps of our recommendation Process

- Relational attributes generation

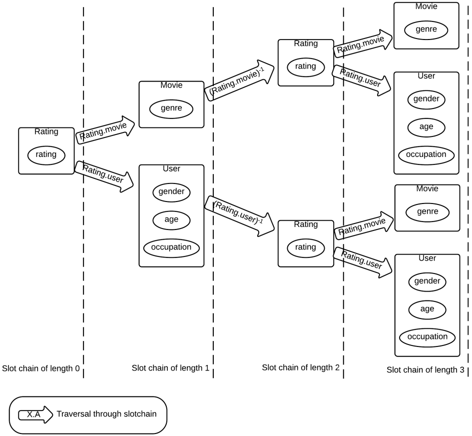


FIGURE : Slot chain traversal

Steps of our recommendation Process

- Relational attributes generation : Multiset operation

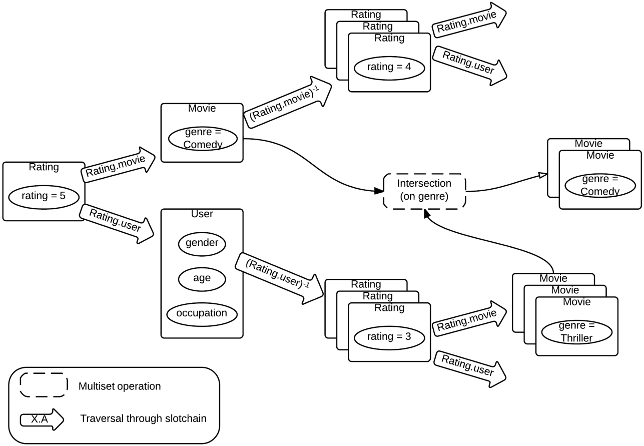


FIGURE : Illustration for multi-set operation

Steps of our recommendation Process

- Markov Blanket detection
  - Identify the attributes within *Markov Blanket* of the target attribute
  - We are using *IAMB (Incremental Association Markov Blanket, [Tsamardinos et al., 2003])* algorithm with *Mutual Information* as the association measure
- Model creation
  - Heuristic search approach can be used to find the exact structure of the attributes
  - We build a Naïve Bayesian classifier
  - Parameters learning
    - Maximum Likelihood Estimation when enough data
    - Experts' knowledge when insufficient data (new system)

Steps of our recommendation Process

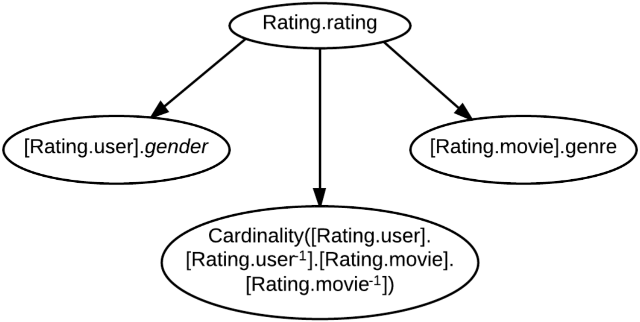


FIGURE : A simple classifier for recommendation

- Recommendation
  - Predict $Pr(V|X)$ where V is the target attribute (Rating.*rating* in our example) and X are the Markov Blanket attributes

Outline...

Incurred difficulties

Existing tools

- Almost there is no software tools for manipulating PRMs :
  - Unbbayes : limited to simple PRM models creation and transformation to GBN. No learning methods have been implemented.
  - Proximity : allow the creation and evaluation of some probabilistic models designed for relational knowledge discovery, not PRMs.
  - RCD/RPC : provide only DAPER learning.

Incurred difficulties

Existing tools

- Almost there is no software tools for manipulating PRMs :
  - Unbbayes : limited to simple PRM models creation and transformation to GBN. No learning methods have been implemented.
  - Proximity : allow the creation and evaluation of some probabilistic models designed for relational knowledge discovery, not PRMs.
  - RCD/RPC : provide only DAPER learning.

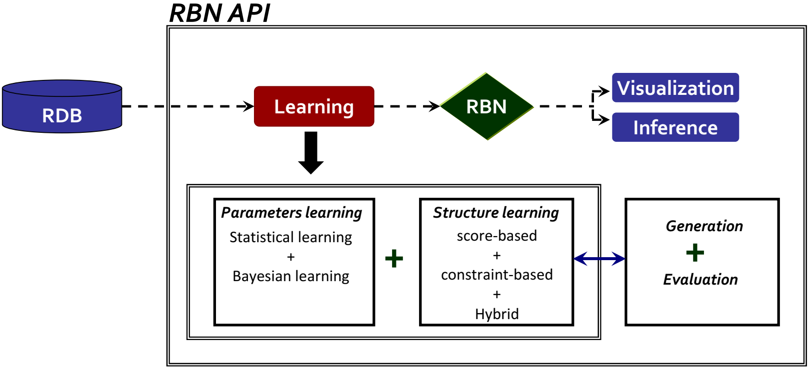
Solution

Need to develop all from scratch.

PILGRIM platform

- Implemented in C++ language
- ProBT API for BN basic creation / parameter learning / inference
- Boost graph library

PILGRIM platform



Outline...

1. Probabilistic relational models
2. Recommendation using PRMs
3. Implementation
4. Conclusion & perspective

Conclusion

- Use PRM to combine simple models – Demographics / Items' features.
- Capable of providing recommendations for new users.
- Can be made capable of providing recommendations in a new system.
- Longer slot chains for collaborative filtering.
- Multiset operations for better attributes.

Perspective

- Use of long slot chains and integration with multiset operators.
- Evaluation on multiple datasets.
- Comparison based on performance metrics.

THANK YOU !

Any question ?

Outline...

Evaluation

- Top-N recommendation method
  - Items are ranked on the basis of the probability of getting high ratings
  - Top N items are recommended and compared with the user's actual preference
  - Average metrics over all users

Evaluation

- Metrics :

$$\text{Precision}(P) = \frac{\text{Number of } \textit{hits}}{N} \tag{1a}$$

$$\text{Recall}(R) = \frac{\text{Number of } \textit{hits}}{\text{Number of high rated movies in the test set}} \tag{1b}$$

$$\text{F-measure}(F) = \frac{2 \times P \times R}{P + R} \tag{1c}$$

where *hits* is the number of actual preferences that appear in the top-N recommendations.

Evaluation

- Dataset – MovieLens dataset
    - 80% train set, 20% test set
    - 2 type of test sets – General test set and test data for coldstart situation

Result

- 2 models
  - *Rating.user.gender*, *Rating.user.age*, *Rating.movie.genre*
  - *Rating.user.gender*, *Rating.user.age*, *Rating.user.occupation*,*Rating.movie.genre*

Test set	Model	N = 10			N = 20		
		P	R	F-M	P	R	F-M
General	Model 1	0.0176	0.0203	0.0186	0.0245	0.0376	0.0279
Coldstart	Model 1	0.0200	0.0200	0.0200	0.0250	0.0268	0.0258
General	Model 2	0.0588	0.0247	0.0213	0.0735	0.0551	0.0393
Coldstart	Model 2	0.0200	0.0200	0.0200	0.0350	0.0350	0.0350