Recommendation Techniques for Multimedia systems

By Sanggil Kang
Department of Computer Science,
The University of Suwon

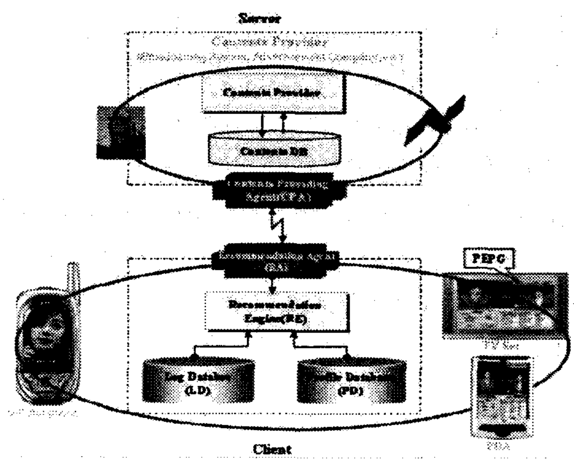
Contents

- Why do we need personalized system in multimedia system?
- Overall Architecture of Multimedia Recommendation System
- Our Developed Recommendation Algorithms
  - Supervised learning of the mutual information
  - Moving average technique
- Conclusion and Future Work

Why do we need personalized system in multimedia system?

- Flood of multimedia content over digital TV channels, the internet, and etc
- Difficulty in finding a user's preferred content
- Spend heavy surfing time
- Likely missing while searching

Overall Architecture of Multimedia Recommendation System



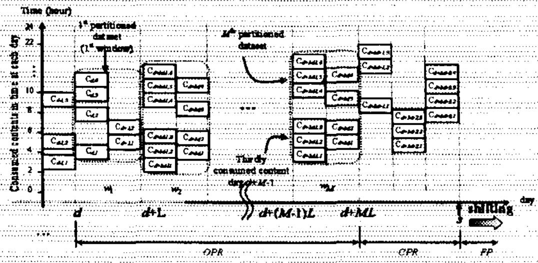
Our Developed Recommendation Algorithms

- Supervised learning of the mutual information

- Moving average technique

Supervised learning of the mutual information

- **Window weights**: providing weights according to the consumed time of the contents



OPR: Old preference region
CPR: Current preference region
FP: Future preference

Supervised learning of the mutual information

- Statistical preference of the i^{th} content x_i

$$\hat{\theta}_{x_i} = p(X = x_i \mid E) = \sum_{m=1}^{M} w_m n_{i,m} / \sum_{m=1}^{M} w_m N_m \qquad (1)$$

Equation (1) depends on the window weights.
→ Accuracy of the preference prediction depends on the window weights

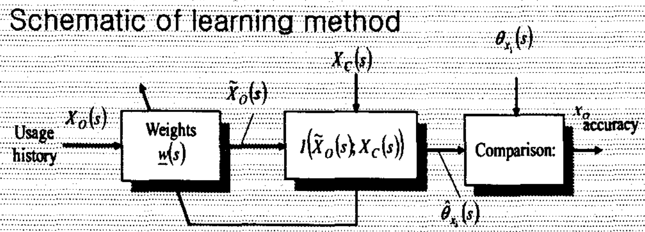
Here,
X : a set of consumed content
N_m: sample number of whole content
$N_{i,m}$: sample number of content x_i within m^{th} window in *OPR*
E: evidence

Supervised learning of the mutual information

- Determination of optimal window weights
  - The latest window give a big impact on predicting the preference
  - The size of the latest window can not be reliable for computing the statistical preference
  - Weights are adjusted in the sense that a set of content in OPR is getting correlated with a set of weighted content in the CPR

Supervised learning of the mutual information

- Schematic of learning method



X_O : *a set of contents in OPR*

X_C : *a set of contents in CPR*

$\tilde{X}_O$: *Weighted version of*

$I(\tilde{X}_O(s); X_C(s))$: *Mutual information between* X_C *and* $\tilde{X}_O$

Supervised learning of the mutual information

- Weight update
  - Mutual information (MI)

$$I(\tilde{X}_O(s); X_C(s)) = \log\left(p(X_C(s) \mid \tilde{X}_O(s)) / p(X_C(s))\right)$$
$$= \log\left(p(X_C(s) \mid \tilde{X}_O(s))\right) - \log\left(p(X_C(s))\right)$$

  - Partial derivate of MI with respective to weight $w_m(s)$

$$\partial I(\tilde{X}_O(s); X_C(s)) / \partial w_m(s)$$
$$= \partial \log\left(p(X_C(s) \mid \tilde{X}_O(s))\right) / \partial w_m(s) - \partial \log\left(p(X_C(s))\right) / \partial w_m(s)$$

Supervised learning of the mutual information

- Amount of update using the delta rule

$$\Delta w_m(e,s) = \eta \cdot \partial I\left(\tilde{X}_O(s), X_C(s)\right) / \partial w_m(s)$$

η : learning rate, e : epoch

- Restriction for weight update

$$\begin{Bmatrix} w_m(e,s) \leftarrow w_m(e-1,s) + \Delta w_m(e,s), & \text{if } w_m(e,s) > 0 \\ w_m(e,s) \leftarrow w_m(e-1,s), & \text{otherwise} \end{Bmatrix} \qquad (2)$$

Experiment (1)

- Prediction of TV viewers' digital TV genre preference
- Experiment conditions
  - **Data**: 2,000 viewers' TV watching history data collected from December 1, 2002 to May 31, 2003, which is provided by AC Nielsen Korea
  - **Evidence**: watching day and time
    E = {(6 p.m. ~ 8 p.m., Monday), (6 p.m. ~ 8 p.m., Monday), . . . , (10 p.m. ~ 12 p.m., Sunday)}
  - **8 genres**: Education, Drama & Movie, News, Sports, Children, Entertainment, Information, Others
  - = 0.1

Experiments (1)

- Computation of prediction error

$$Error(s) = \sum_{i=1}^{J} \left| \hat{\theta}_{x_i}(s) - \theta_{x_i}(s) \right| \qquad (3)$$

Here,

$\theta_{x_i}(s)$: True preference of x_i at shift s

$\hat{\theta}_{x_i}(s)$: Estimated preference of x_i at shift s

Experiments (1)

- Experimental results

Age (Gender)	Method	The number of windows in the OPR, M					
		3	4	5	6	7	8
10s (male)	Typical	0.24	0.21	0.22	0.2	0.24	0.23
	Our	0.12	**0.1**	0.13	0.14	0.16	0.15
10s (female)	Typical	0.22	0.2	0.21	0.24	0.23	0.25
	Our	**0.13**	0.15	0.16	0.15	0.14	0.15
20s & (male)	Typical	0.22	0.21	0.19	0.21	0.2	0.22
	Our	0.14	0.15	**0.12**	0.14	0.16	0.17
20s & female	Typical	0.2	0.19	0.23	0.19	0.21	0.22
	Our	0.14	**0.11**	0.13	0.15	0.13	0.12
30s & (male)	Typical	0.22	0.23	0.2	0.22	0.22	0.23
	Our	0.13	0.14	**0.12**	0.13	0.14	0.13
30s &(female)	Typical	0.21	0.19	0.2	0.19	0.22	0.22
	Our	0.14	0.15	0.11	**0.11**	0.15	0.13
40s & (male)	Typical	0.23	0.22	0.23	0.21	0.19	0.22
	Our	**0.16**	0.17	0.17	0.17	**0.16**	0.18
40s& (female)	Typical	0.19	0.18	0.19	0.18	0.2	0.19
	Our	0.17	**0.16**	0.17	**0.16**	0.17	0.18
50s & (male)	Typical	0.14	0.15	0.14	0.15	0.14	0.14
	Our	0.15	0.14	0.13	0.13	0.13	**0.12**
50s& (female)	Typical	0.13	0.14	0.15	0.14	0.14	0.15
	Our	0.15	0.13	0.15	0.13	**0.12**	**0.12**

Moving Average Technique

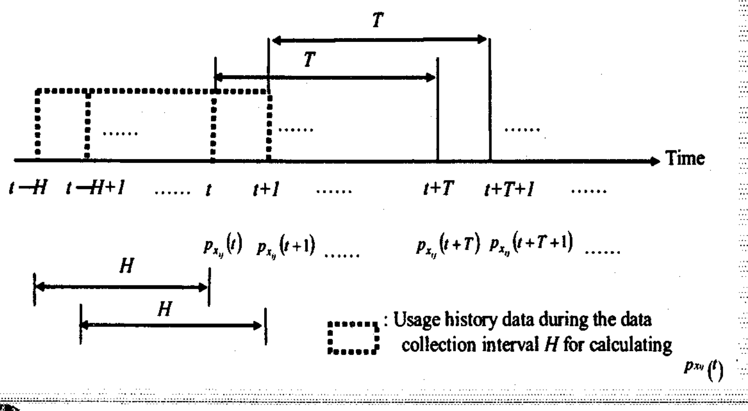
- Drawback of existing preference techniques:

  Not suitable for representing the dynamically changing users' preference because the history of users' preference change is not considered for updating the statistical preference
- Moving Average Technique:

  Considering the variations of the past preference changes by adding the mean variation of the past statistical preference to the current statistical preference of content

Moving Average Technique

- Schematic representation



Moving Average Technique

- Estimation of statistical preference
  - The statistical preference of content x_i at time t

$$p_{x_i}(t) \equiv p(x_i \mid E, t, H) = \frac{n_{i,H}(t)}{N_H(t)} \qquad (4)$$

Here, $N_H(t)$: Total sample number at time t collected with H (window size)
$n_{i,H}(t)$: Sample number of content x_i at time t with H (window size)

  - Variance of Statistical preference (VSP) between two consecutive statistical preferences

$$\Delta p_{x_{ij}}(t+1) = p_{x_{ij}}(t+1) - p_{x_{ij}}(t)$$

Moving Average Technique

- A set of VSPs during a predetermined period T from t to $t+T$

$$\Delta p_{x_{ij}}(t, t+1) = [\Delta p_{x_{ij}}(t+1), \Delta p(t+2), \cdots, \Delta p_{x_{ij}}(t+T)]$$

- Update of the statistical preference by adding mean of the set of VSPs ($\Delta \bar{p}_{x_{ij}}(t, t+T)$)

$$\hat{p}_{x_{ij}}(t+T+1) = p_{x_{ij}}(t+T) + \Delta \bar{p}_{x_{ij}}(t, t+T) \qquad (5)$$

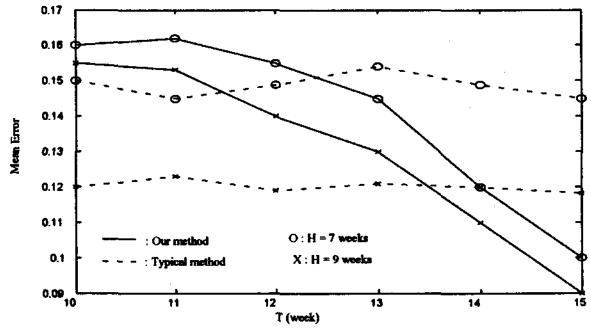
Here, $\Delta \bar{p}_{x_{ij}}(t, t+T) = (1/T)\sum_{k=1}^{T} \Delta p_{x_{ij}}(t+k)$

- Typical method

$$\hat{p}_{x_{ij}}(t+T+1) = p_{x_{ij}}(t+T) \qquad (6)$$

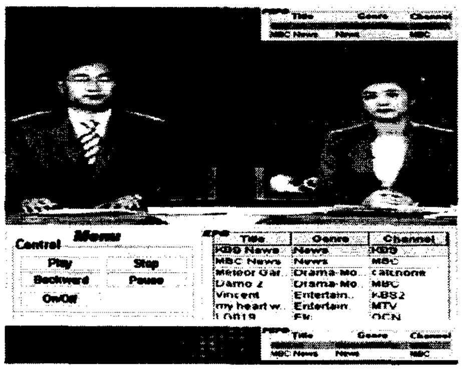
Experiment (2)

- Experimental results with same data used in the supervised learning method



Personalized Program Guide

- An example of personalized program guide displaying a TV user's preferred TV programs inferred by our methods



Conclusion and Future Work

- Conclusion
  - Our algorithms outperformed typical methods
  - Adaptation of the change of the trend of users' preference
- Future Work
  - TV program recommendation using multi-recommendation systems
    - Genre recommendation
    - Channel recommendation
  - Redo experiments using more data
  - Develop methods for determining optimal parameters
  - Apply our method to mobile environment