Upper Bounds For the Correlated Bayesian Information Filtering Problem

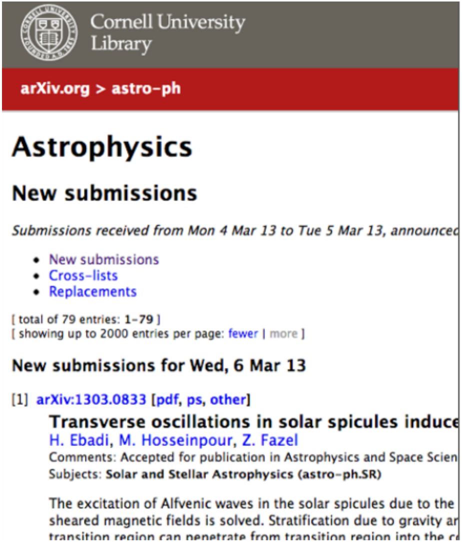
2015 Applied Probability Conference

Bangrui Chen, Peter Frazier

Cornell University
Operations Research and Information Engineering
bc496@cornell.com, pf98@cornell.edu

July 6th, 2015

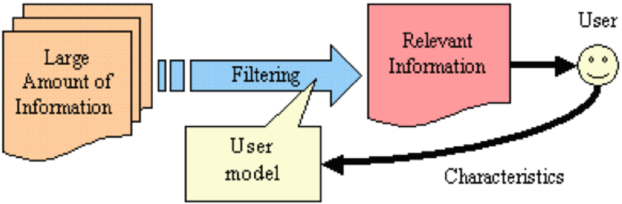
Information Filtering



- There are lots of new papers (roughly 80 new papers/day in astrophysics.
- Problem 1: Browsing this many papers is a lot of work for researchers.
- Problem 2: Researchers still miss important developments.

Information Filtering

- Information filtering systems automatically distinguish relevant from irrelevant items (emails, news articles, intelligence information) in large information streams.



Exploration vs. Exploitation

- "cold-start": sufficient training data may not be available.

Exploration vs. Exploitation

- "cold-start": sufficient training data may not be available.
- Exploration vs. exploitation tradeoff.

Traditional Approach

- Does not actively explore to get the most useful feedback, such as:
    - Exploration at the beginning followed by exploitation.
    - Exploration and exploitation periodically.

Traditional Approach

- Does not actively explore to get the most useful feedback, such as:
    - Exploration at the beginning followed by exploitation.
    - Exploration and exploitation periodically.
- We present a Bayesian sequential decision-making formulation of this problem.

Our Results

- An instance-specific computational upper bound on the value of a Bayes-optimal strategy.

Our Results

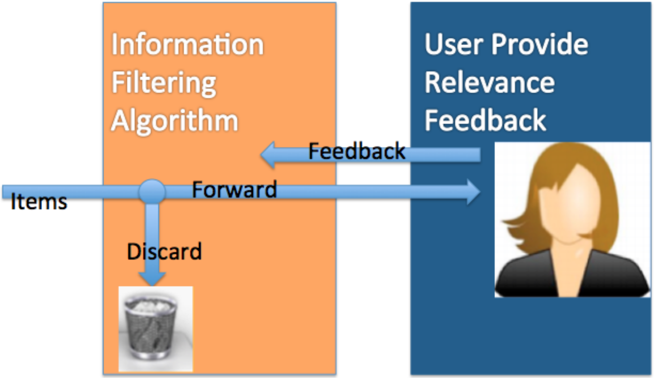
- An instance-specific computational upper bound on the value of a Bayes-optimal strategy.
- An implementable policy, called Decompose Then Decide (DTD).

Our Results

- An instance-specific computational upper bound on the value of a Bayes-optimal strategy.
- An implementable policy, called Decompose Then Decide (DTD).
- Prove our upper bound is asymptotically tight.

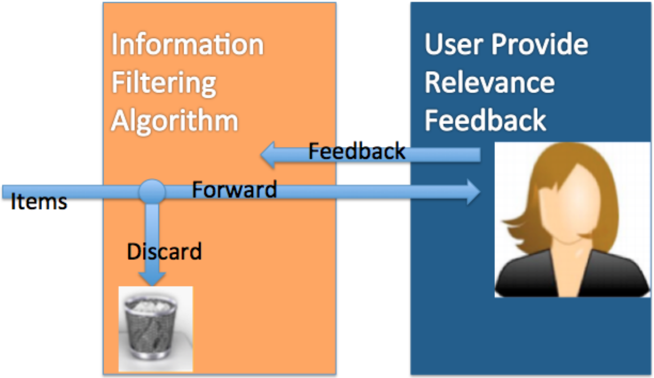
Bayesian Model

- We consider information filtering for a single user.



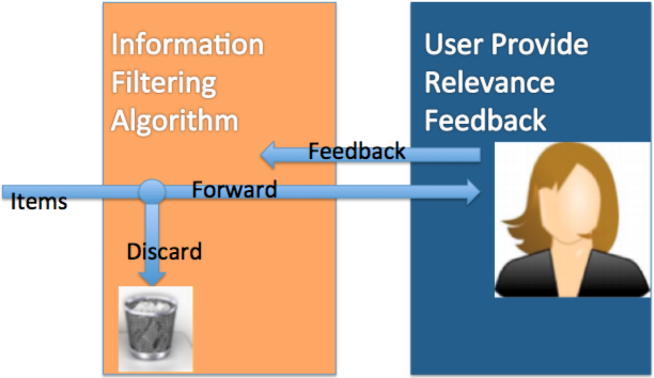
Bayesian Model

- We consider information filtering for a single user.
- The n^{th} arriving item is described by a k-dimensional feature vector $X_n = \{x_{1,n}, \cdots, x_{k,n}\}$. We assume that $x_{i,n} \geq 0$ for all i and n and $\sum_{i=1}^{k} x_{i,n} = 1$.



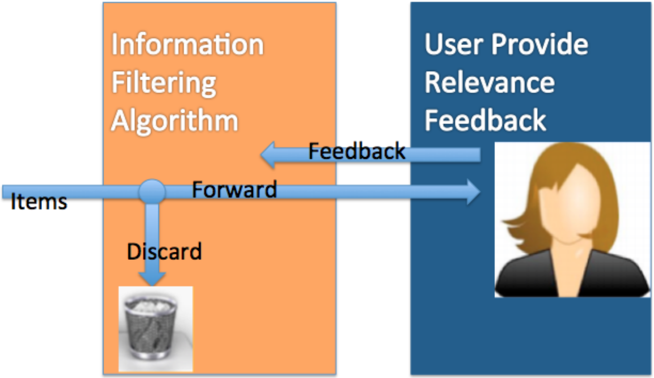
Bayesian Model

- Let $U_n \in \{0,1\}$ represent this decision for the n^{th} item, where 1 means to forward and 0 means not to forward.



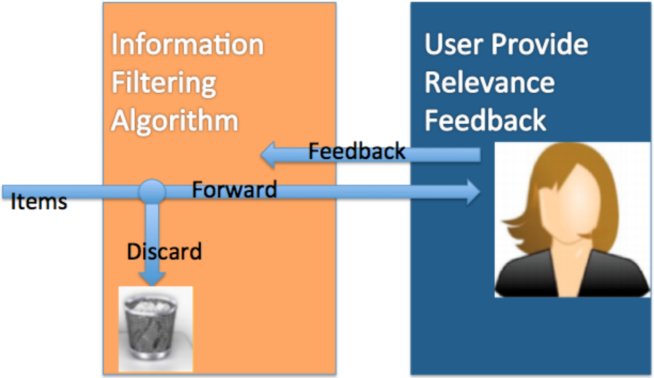
Bayesian Model

- Let $U_n \in \{0,1\}$ represent this decision for the n^{th} item, where 1 means to forward and 0 means not to forward.
- Each time the system forwards, it pays a constant cost c and receives the item's relevance Y_n as a reward.



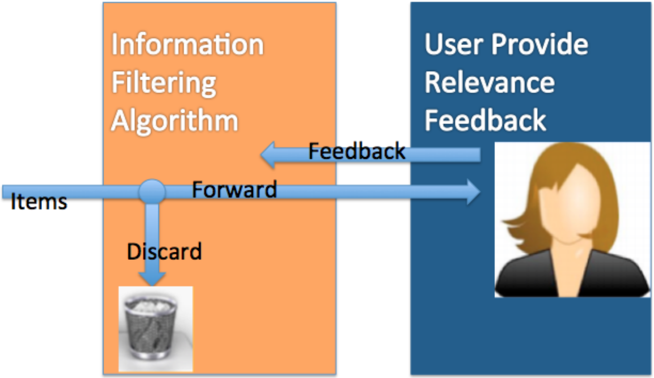
Bayesian Model

- Let $\theta = (\theta_1, \cdots, \theta_k)$ denote the single user's latent preference vector for the k different features.



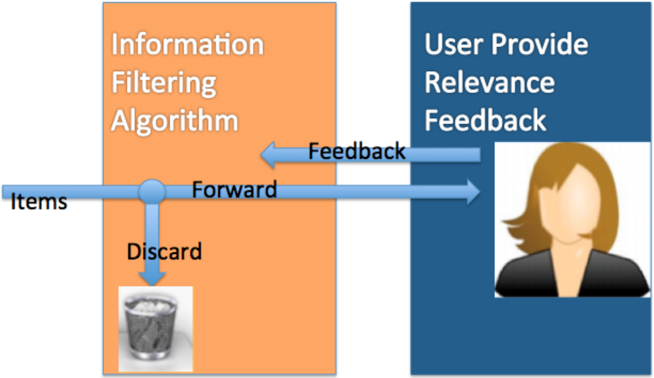
Bayesian Model

- Let $\theta = (\theta_1, \cdots, \theta_k)$ denote the single user's latent preference vector for the k different features.
- We assume $\theta \sim N(\mu_0, \Sigma_0)$.



Bayesian Model

- Let $\theta = (\theta_1, \cdots, \theta_k)$ denote the single user's latent preference vector for the k different features.
- We assume $\theta \sim N(\mu_0, \Sigma_0)$.
- This relevance is modeled as $Y_n = \theta \cdot X_n + \epsilon_n$, where $\epsilon_n \sim N(0, \lambda^2)$ (Similar in spirit to Bayesian linear bandit).



Our Goal

- The decision of whether or not to forward the n^{th} item can only depend on our current X_n as well as the previous information $H_{n-1} = (U_m, X_m, U_m Y_m : m \leq n-1)$.

Our Goal

- The decision of whether or not to forward the n^{th} item can only depend on our current X_n as well as the previous information $H_{n-1} = (U_m, X_m, U_m Y_m : m \leq n-1)$.
- Mximize the following:

$$\sup_{\pi \in \Pi} E^{\pi}\left[\sum_{n=1}^{N} U_n(Y_n - c)\right] \tag{1}$$

where E^{π} denotes the expected reward using policy π.

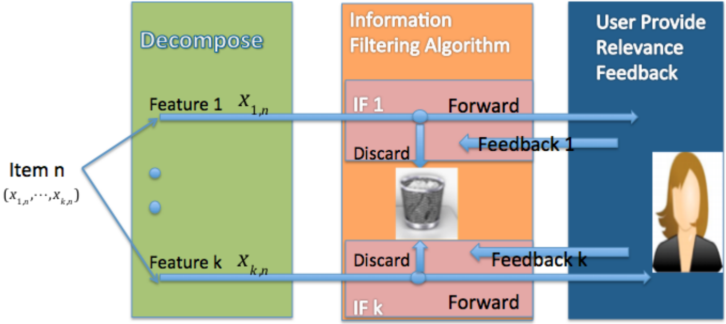
1. Motivation
2. Problem Formulation
3. Main Theoretical Results
4. Numerical Experiments
5. Conclusion

Upper Bounds

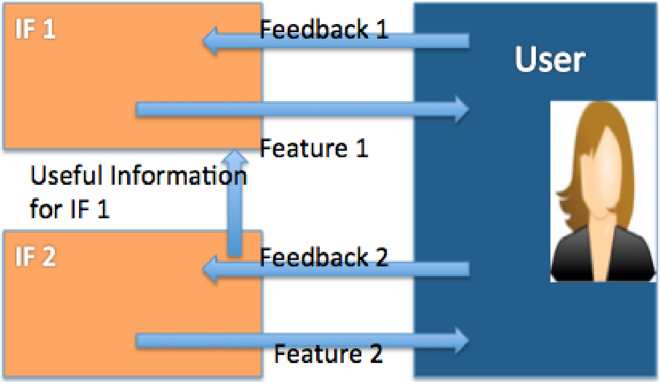
- This problem can be solved using stochastic dynamic programming, but it suffers from curse of dimensionality.

Upper Bounds

- This problem can be solved using stochastic dynamic programming, but it suffers from curse of dimensionality.
- Our idea is:



Why we need information relaxation?



The maximum information Feedback2 can contain is θ_2, which corresponds to the user's latent preference vector for the second feature.

Upper Bounds

- Define $Y_{i,n} = \theta_i + \epsilon_n^i$ and $\epsilon_n^i \sim N(0, \lambda^2)$ for $i = 1, 2, \cdots, k$. We may think of $Y_{i,n}$ as the reward that we would have seen if X_n were equal to e_i.

Upper Bounds

- Define $Y_{i,n} = \theta_i + \epsilon_n^i$ and $\epsilon_n^i \sim N(0, \lambda^2)$ for $i = 1, 2, \cdots, k$. We may think of $Y_{i,n}$ as the reward that we would have seen if X_n were equal to e_i.
- Rewrite $Y_n = \sum_{i=1}^k x_i\theta_i + \epsilon_n = \sum_{i=1}^k x_i(\theta_i + \epsilon_n^i) = \sum_{i=1}^k x_i Y_{i,n}$.

Upper Bounds

- Define $U_{j,n}$ to be decision made for the j^{th} feature of the n^{th} item.

Upper Bounds

- Define $U_{j,n}$ to be decision made for the j^{th} feature of the n^{th} item.
- For each feature j, we now introduce a new set of policies Π_j and the decision in Π_j may depend on $\theta \cdot e_i$ for $\forall i \neq j$.

Upper Bounds

Theorem 1

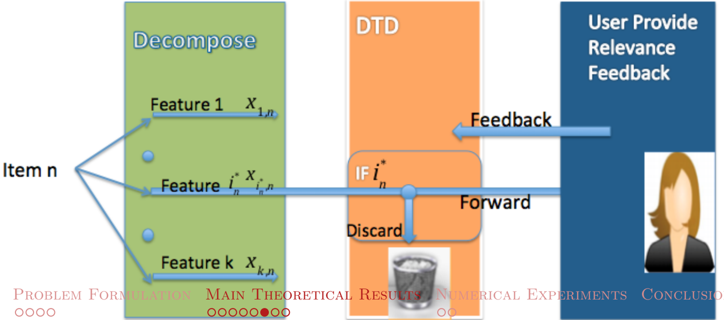
$$\sup_{\pi \in \Pi} E^{\pi}\left[\sum_{n=1}^{N} U_n(Y_n - c)\right] \leq \sum_{j=1}^{k} \sup_{\pi'' \in \Pi_j} E^{\pi''}\left[\sum_{n=1}^{N} U_{j,n}(x_{j,n} Y_{j,n} - x_{j,n} c)\right].$$

Decompose Then Decide (DTD)

DTD

Let π_i^* denote the optimal policy for the i^{th} single-feature subproblem. Then, DTD is defined as follows:

1. Find $i_n^* \in \arg\max\{X_n e_i, i = 1, \cdots, k\}$.
2. Set $U_n = \pi_{i_n^*}^*(\mu_{i_n^*}, \sigma_{i_n^*})$.
3. Update μ, Σ respectively.



Asymptotic optimality

Let $(\mathbb{P}^l : l = 1, 2, \cdots)$ be a sequence of probability distribution on R^k with $\mathbb{P}^l \to \mathbb{P}_*$ in distribution, where $\mathbb{P}_*$ has support on the unit vectors.

- Let $U(\mathbb{P})$ denote the upper bound from Theorem 1 in which $X_n \sim \mathbb{P}$.
- Let $V(\mathbb{P})$ denote the value of an optimal policy (i.e., the value function), and $V^{DTD}(\mathbb{P})$ as the value of the DTD policy, for this same problem instance.

Asymptotic optimality

Theorem 2

Suppose that Σ_0 is a diagonal matrix. Then

$$\lim_{l\to\infty} (U(\mathbb{P}^l) - V(\mathbb{P}^l)) = 0,$$

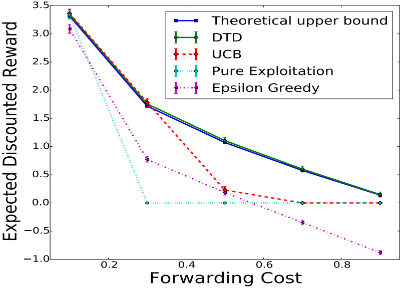
$$\lim_{l\to\infty} V(\mathbb{P}^l) - V^{DTD}(\mathbb{P}^l) = 0,$$

$$\lim_{l\to\infty} V(\mathbb{P}^l) = U(\mathbb{P}_*) = V(\mathbb{P}_*).$$

Numerical Experiments

- Simulated Data.
- Yelp Academic Dataset.

Numerical Experiments



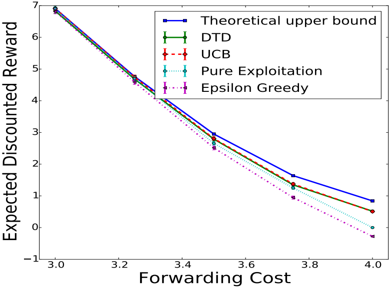


Figure : The left plot compares performance on simulated data and the right plot compares performance on the Yelp academic dataset.

1. Motivation
2. Problem Formulation
3. Main Theoretical Results
4. Numerical Experiments
5. Conclusion

Conclusion

- An instance-specific computational upper bound.
- An implementable policy (DTD).
- Prove asymptotically optimality.

Question?

Thanks for your time!