Information Filtering for arXiv.org:

Bandits, Exploration vs. Exploitation, and the Cold Start Problem

Peter Frazier, Xiaoting Zhao
School of Operations Research & Information Engineering
Cornell University

Fusion Fest, DIMACS, Rutgers University, October 11th 2014
Supported by NSF BIGDATA 1247696

This work is part of an NSF grant with Paul

Paul Kantor (PI)

Dave Blei

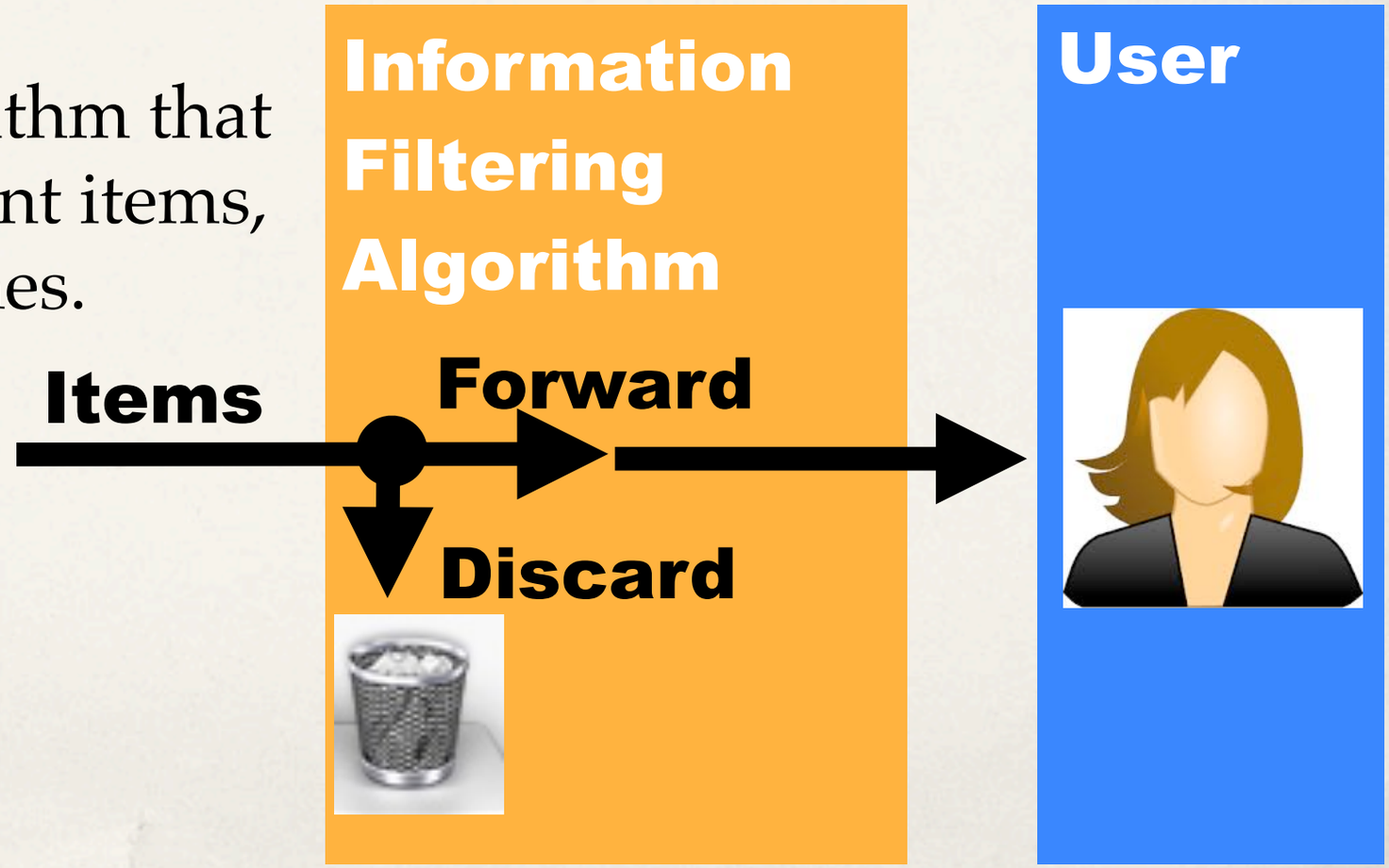
Paul Ginsparg

Thorsten Joachims

We are interested in

information filtering

- We face a sequence of time-sensitive items (emails, blog posts, news articles).
- A human is interested in some of these items.
- But, the stream is too voluminous for her to look at all of them.
- We wish to design an algorithm that forwards most of the relevant items, and few of the irrelevant ones.



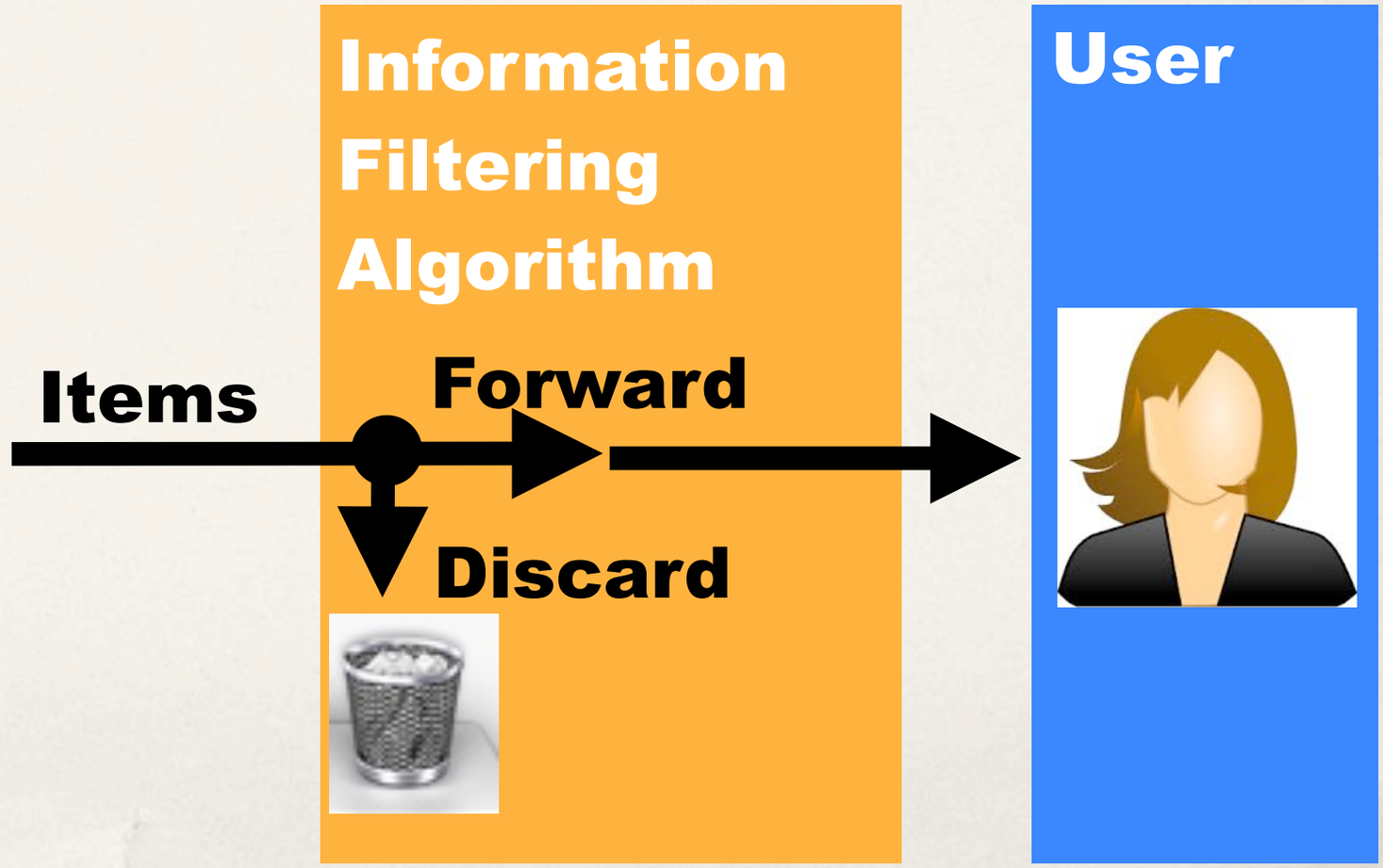
We are interested in

information filtering

- If we had lots of historical data, we could train a machine learning classifier to predict which items would be relevant to this user.

- But what if we are doing information filtering for a new user, i.e., from a **cold start**?

- How can we quickly learn user preferences, without forwarding too many irrelevant items?



We are interested in

exploration vs. exploitation

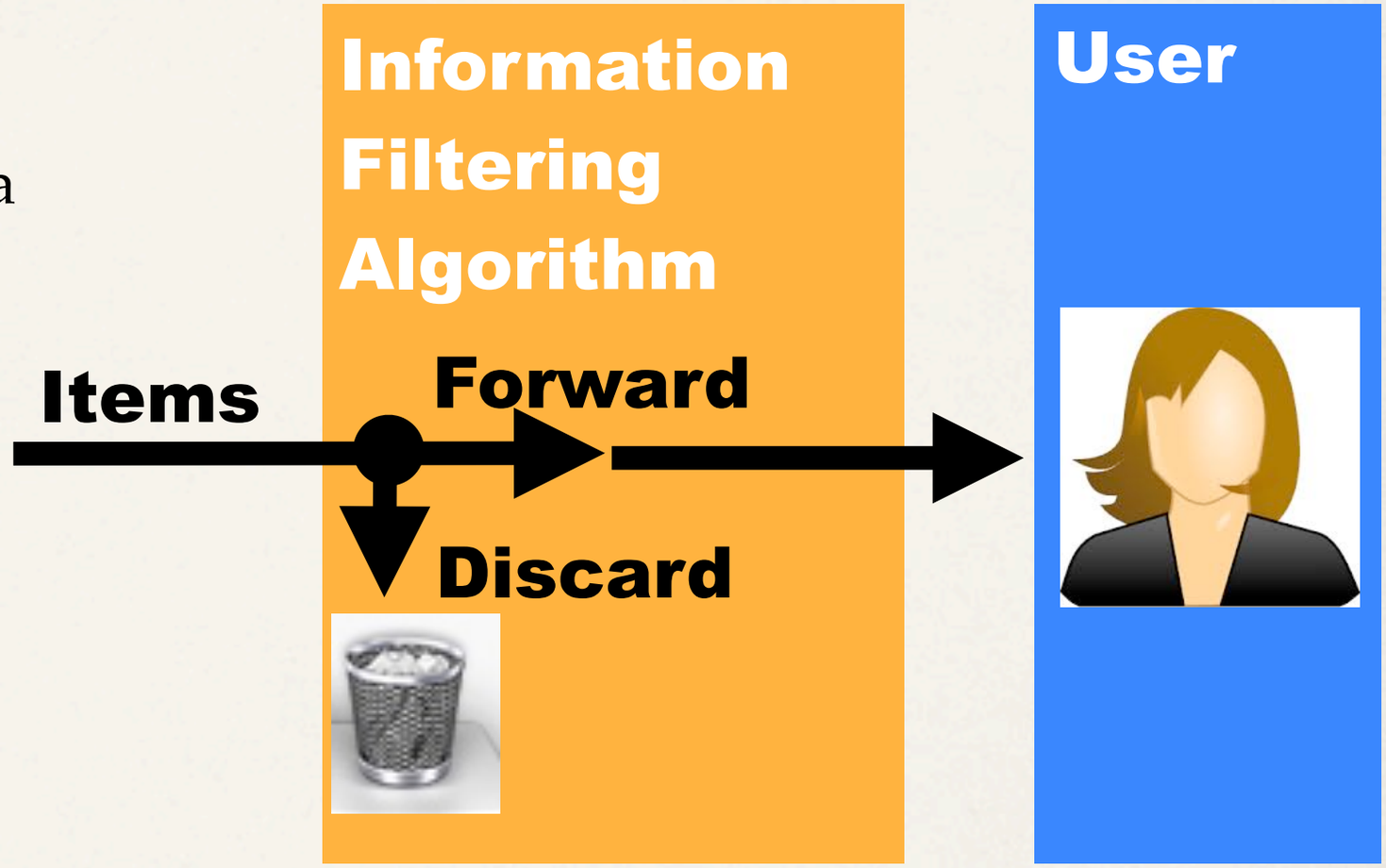
in information filtering

- What if we are filtering for a new user, or filtering items of a type we haven't seen before?

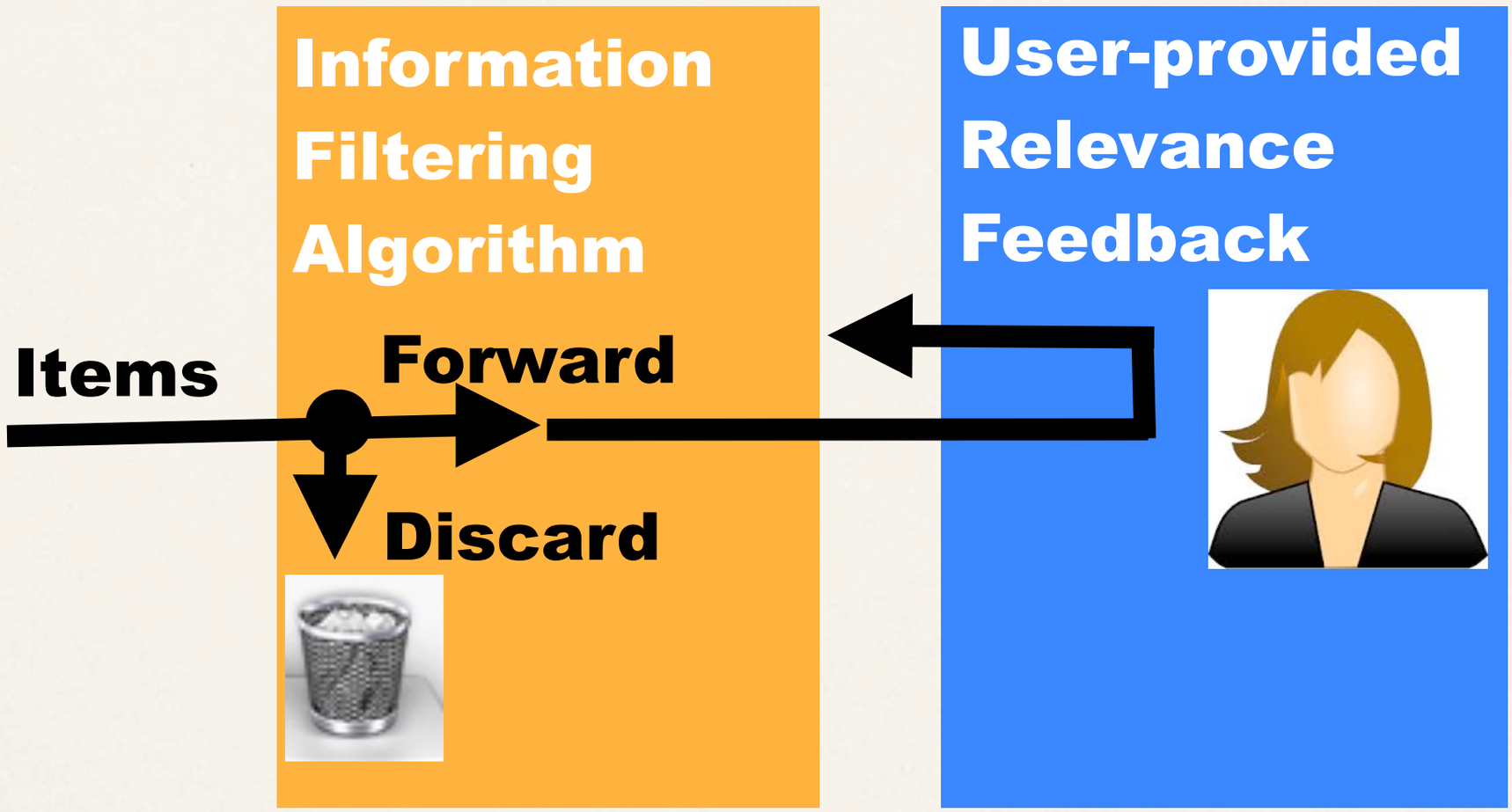
- We may want to **EXPLORE**, i.e., forward a few items of unknown relevance, to allow learning.

- But, we may want to **EXPLOIT** what little training data we have, which may suggest these items type is irrelevant.

- What should we do?

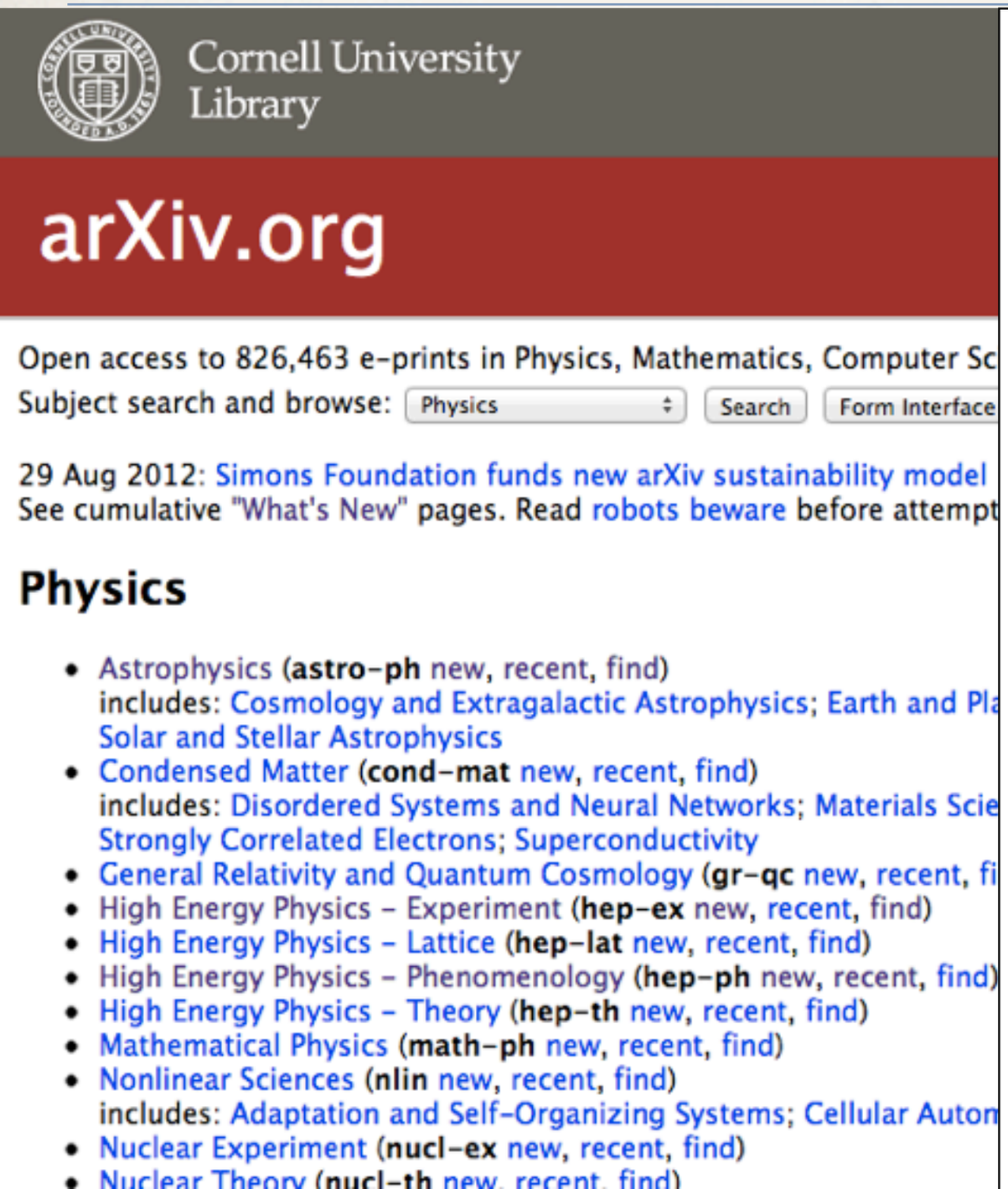


We develop an information filtering algorithm that trades exploration vs. exploitation



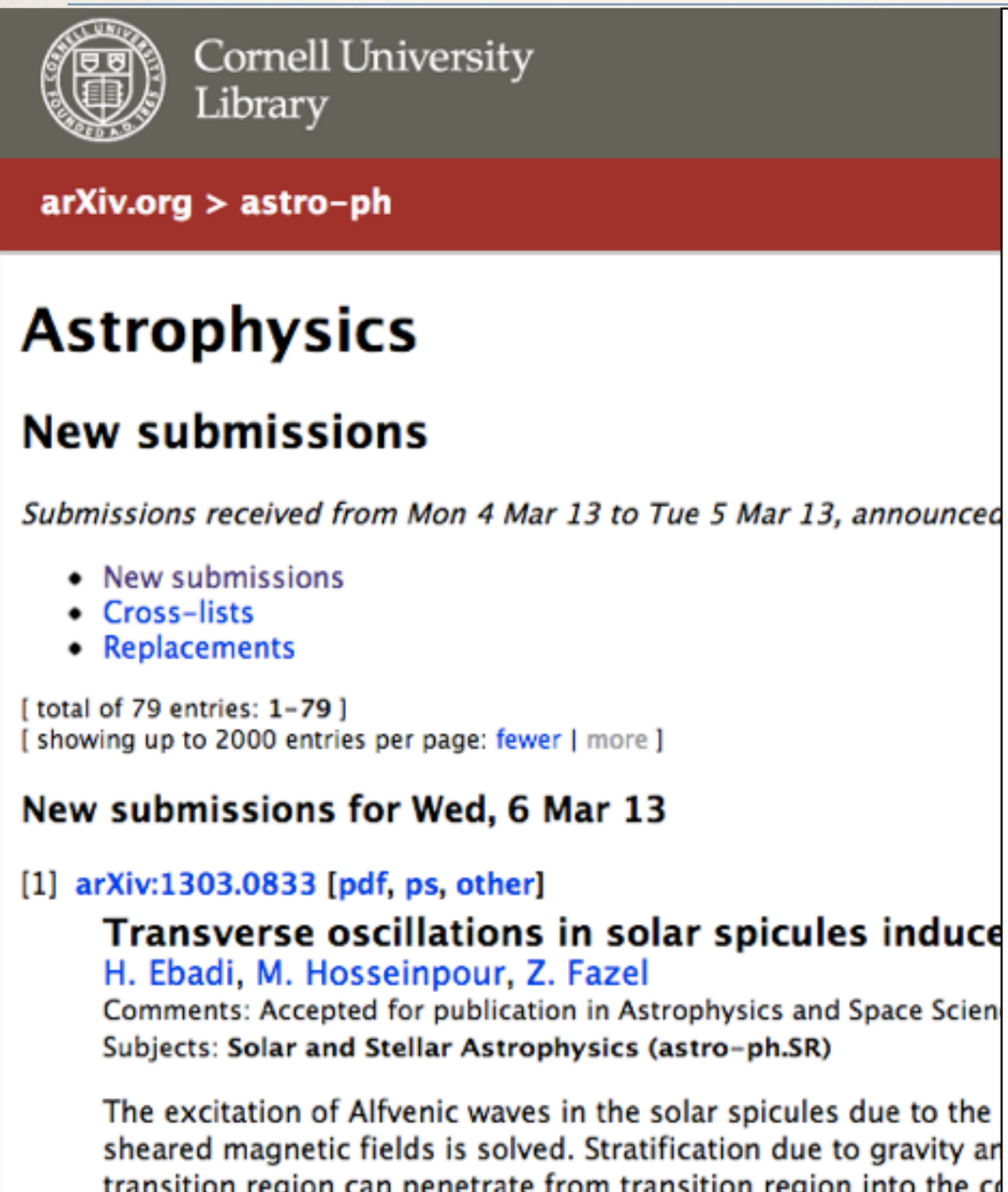
- We use **dynamic programming** and a **Bayesian analysis** to provide an algorithm that is **average-case optimal** for a particular version of the information filtering problem.

We are motivated by an information filtering system we are building for arxiv.org



- arXiv.org is an electronic repository of scientific papers hosted by Cornell.
- Papers are in physics, math, CS, statistics, finance, and biology.
- arXiv currently has ≈800,000 articles, and 16 million unique users accessing the site each month.

Our goal is to improve daily & weekly new-article feeds



- Many physicists visit the arXiv every day to browse the list of new papers, to stay aware of the latest research.
- There are lots of new papers: e.g., 15 new papers / day in arxiv category astro.GA, "Astrophysics of Galaxies."
- Problem 1: Browsing this many papers is a lot of work for researchers.
- Problem 2: Researchers still miss important developments.

Literature Review

- Exploration vs. exploitation has been studied extensively in the multi-armed bandit problem:
  - Bayesian treatments: [Gittins & Jones, 1974; Whittle 1980] ...
  - non-Bayesian treatments: [Auer, Cesa-Bianchi, Freund, Schapire, 1995; Auer, Cesa-Bianchi & Fischer, 2002] ...
- Exploration vs. exploitation has been studied in information retrieval: [Zhang, Xu & Callan 2003; Agarwal, Chen & Elango 2009; Yue, Broder, Kleinberg & Joachims 2009; Hofmann, Whitestone & Rijke 2012]

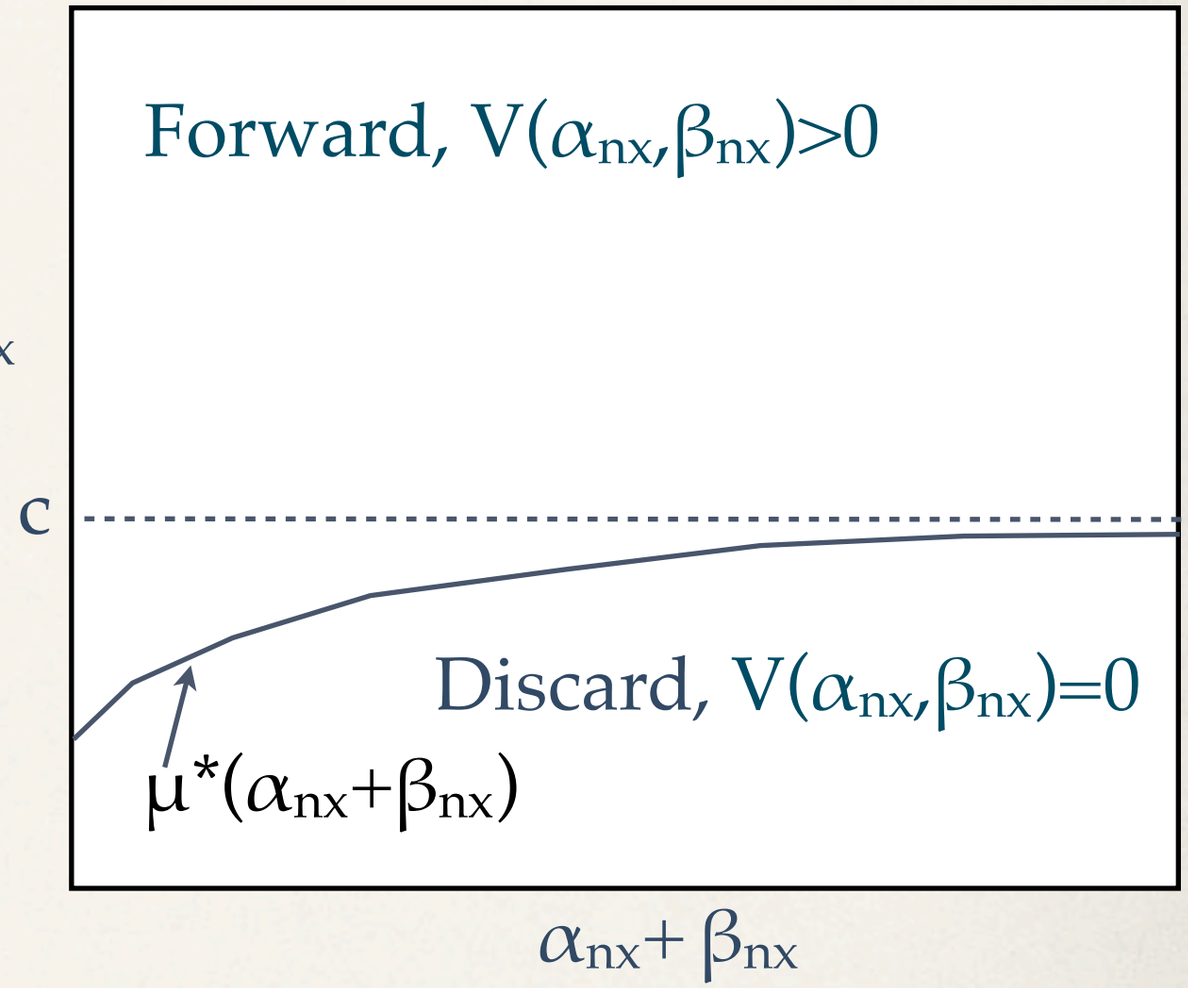
I'll use a simple model to explain the main idea.

- Items are pre-categorized into one of k categories, and the category is the only information about them we use.
- Items within category x are relevant with probability θ_x.
- θ_x is unknown, but we have a Beta(α_{0x}, β_{0x}) prior on it, learned from historical data.
- We only observe relevance of forwarded items. [So **the only way to learn is to forward**.]
- For each forwarded item, we get a reward of 1-c if it is relevant, and pay a penalty of -c if it is irrelevant.
- The user spends a random geometrically-distributed amount of time using our system.
- We wish to maximize expected total reward over the user's time using our system.

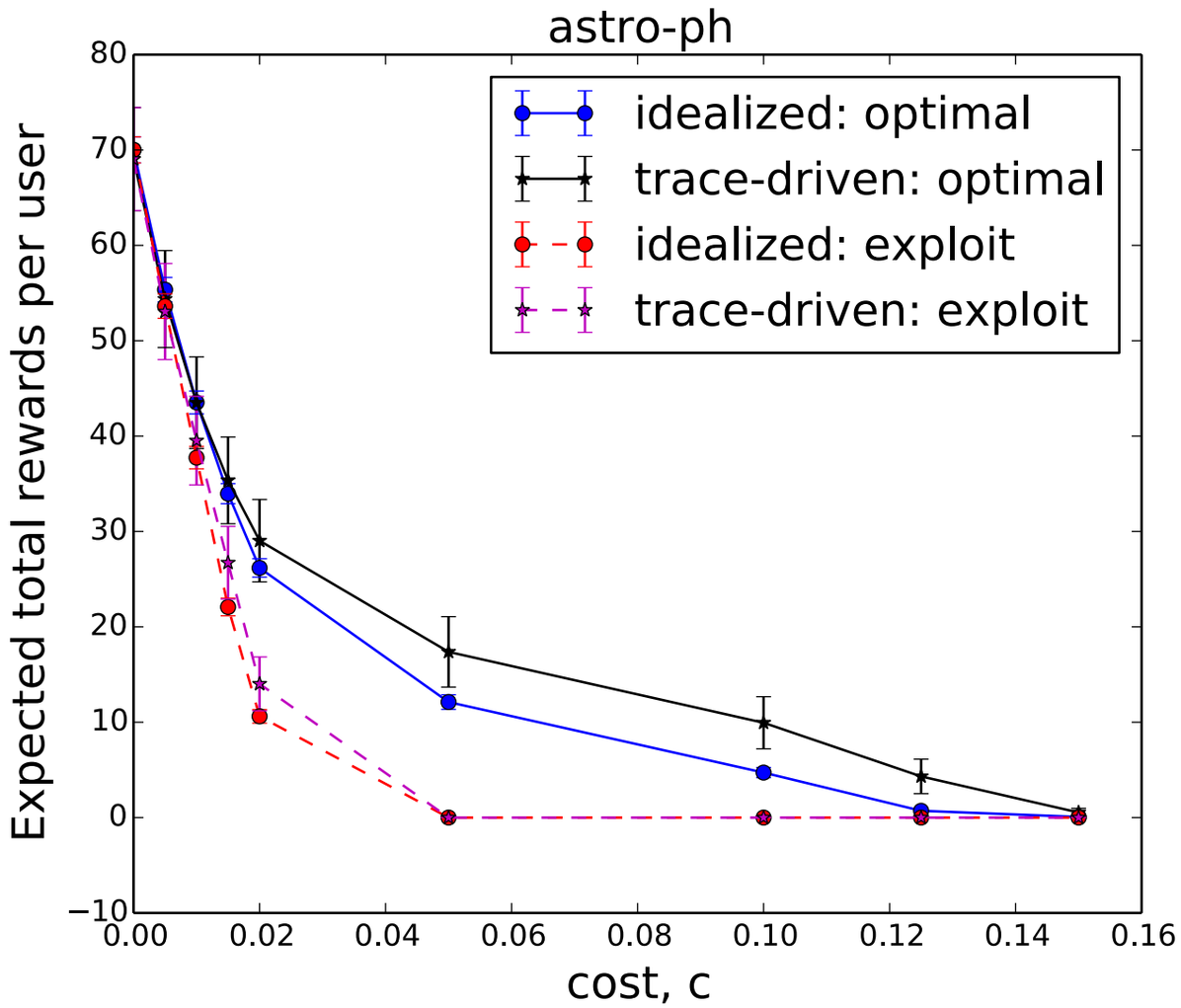
The optimal algorithm looks like this, and can be computed using stochastic dynamic programming.

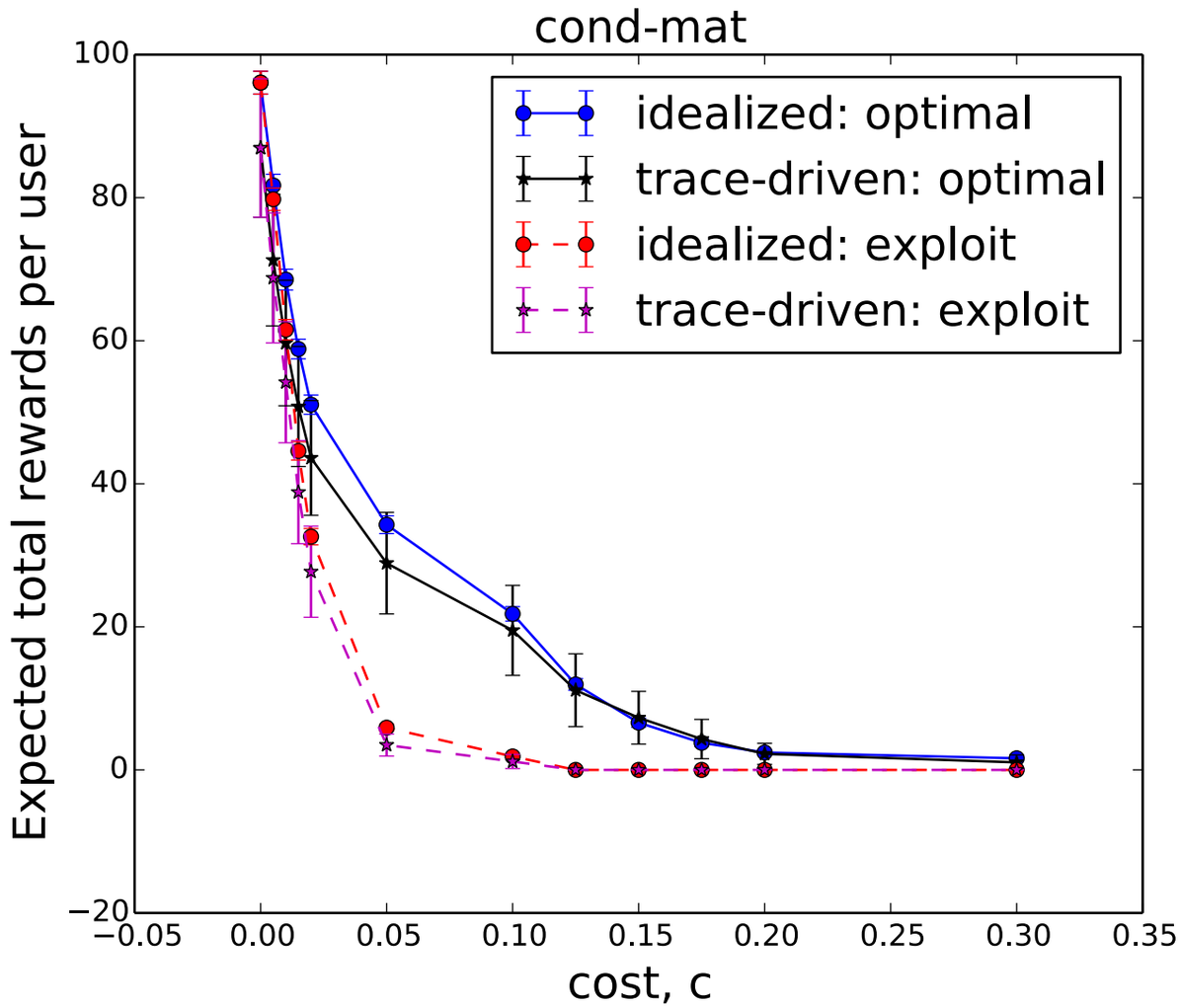
- **Theorem 1**: There exists a function $\mu^*(\cdot)$ such that it is optimal to forward when $\mu_{nx} \geq \mu^*(\alpha_{nx}+\beta_{nx})$ and to discard otherwise.

- **Theorem 2**: $\mu^*(\alpha+\beta)$ has the following properties:
  - it is bounded above by c;
  - it is increasing in $\alpha+\beta$;
  - it goes to c as $\alpha+\beta \to \infty$.



Optimal outperforms myopic in the multi-category problem, in idealized and trace-driven simulations.





We build on this analysis to study more complex models

- **Periodic review**: If the user responds to forwarded items not immediately but only periodically when visiting our website, then our decision is the # of items from each category to show.

- **Rankings**: If the user does not tell us the cost of his time c, and instead examines papers from a ranked list on each visit until his "patience budget" is exhausted, then we can view c as a Lagrange multiplier, and use our analysis to provide a ranking. [Analysis gives an upper bound on the value of the Bayes-optimal procedure.]

- **Linear models**: If items are described by feature vectors rather than categories, and user preference is described by a linear model, then upper bounds on the Bayes-optimal procedure may be derived.

Conclusion

- We presented an information filtering problem arising in the design of a recommender system for arXiv.org
  - We gave details of a simple model, which assumed a known cost, and instantaneous feedback from the user.
  - This model can be extended to periodic review, in which the user provides feedback on items in batches, and to provide rankings over items.
- We are in the process of testing this system, and rolling it out to users of the arXiv.