CS6604
Recommender Systems

Department of Computer Science
Virginia Tech, VA 24061
Spring 2001

Who? Where? When?

- **Instructor**
  - Naren Ramakrishnan, 629 McBryde, 231-8451, naren@cs.vt.edu
  - Office Hours: MW 2pm-4pm (or WIAT)
- **Class Meeting Times**
  - MWF 12:20-1:10pm, McBryde 322
- **Keeping in Touch**
  - Web Page: [search for it!]
  - Listserv: cs6604_15091@listserv.vt.edu
- **Course Format**
  - 1st two weeks: primarily lectures
  - Later: short 25-30 min. lectures + discussions led by instructor
- **Pre-requisites**
  - [CS5114, CS5485, CS5604, CS5614, CS5724, CS5714]$^{+}$
  - Ability to search for and find something on the Internet

What are recommender systems by the way?

- **Definitions will come last!**
  - Too many threads to summarize meaningfully at this point!
- **Various Viewpoints**
  - (the what) customize information access
  - (the why) reduce information overload
  - (the so-what?) retain customers, increase $$$, and other jollies
- **CS 6604 adopts a broader view**
  - includes personalization, HCI models, system design...
- **... in addition to**
  - data mining, web engines, graph theory, and numerical analysis

Examples of Systems

- **Everyday search engines**
  - “Public Transportation”
- **Book Recommenders**
  - amazon.com, bn.com etc.
- **Social Network Navigation (finding experts)**
  - Corporate intranets, “expert sites”, VTED
- **Adaptive Web Sites**
  - weather pages, movie-listings (movies.com), integrated one-stop sites
- **“My”pages**
  - myvt.edu, mycnn.com, myyahoo.com etc.
- **Product placement (Niche finding)**
  - vaio.com, IBM Thinkpads
- **Mobile Access**
  - proxy agents, news-on-demand etc.

Recommender Systems (Contd.)

- **Why study this area?**
  - Excellent integrated view of CS research issues
  - Highly relevant in today's personalized world
  - Experience first-hand a developing research field
- **What to expect in CS 6604**
  - Critical ability to review and evaluate research
  - Digest 5+ research papers a week
  - A project that is your ticket to fame
  - Sleepless nights
- **Among the things you will learn**
  - "It's not technical, s****d!"
  - Not all published work is good (or even correct)
  - Every research area has its share of quacks
  - Good research requires asking tough questions (of yourself and the field)
  - Cartesian product research rarely leads to insights

Grading etc.

- **25%: Class Participation**
  - Devour readings beforehand
  - Contribute (constructively) to the topic
  - Give opportunities for others
  - Every student writes a critical 2-3 page review of a paper
  - (Opt): Volunteer as scribes for summarizing discussion
- **75%: Semester Project**
  - Intensive study/work in an area/topic of interest
  - Review current literature
  - Produce publish-quality report (content as well as presentation)
  - Submit to conferences and journals
  - Ideas for projects are posted on the class pages
- **"How do we know if our project can lead to publication?"**
  - "I will work with you to define and delineate the scope."
  - Try to choose a topic that dovetails with your research interests/program
  - Cross-disciplinary topics most often lead to interesting research!

Ground Rules

- **For reviewing papers**
  - Apply onion principle
  - Always question assumptions
  - Don't get intimidated by theorems, lemmas, and buzzwords
  - Form your own mental model of the area
- **For your project**
  - Do literature survey (web search)
    - *it has probably been done before*
  - explore connections (even if they look tenuous)
  - Are you solving the right problems?
  - Be diligent about experimental evaluation and interpreting results

Recommender Dichotomies

- **How to model a recommender system?**
  - Content-Based: Features
  - Collaborative: Ratings, Evaluations, Experiences
  - Hybrid: Mix of both
- **How to model a recommender system? (alt.)**
  - Weak-theory approach
  - Strong-theory approach
- **How to build a recommender system?**
  - "Public Transportation"
  - "Hot Rods"
- **How to target a recommender system?**
  - information tailored per individual
  - information targeted for clusters of users
  - information useful for everybody (top N lists)
- **How to maintain a recommender system?**
  - offline ("frozen designs")
  - online (incorporate new data on a continual basis)