Research Summary

Anand Bhaskar

I am working on possible heuristics for spam detection. More specifically, I am currently determining how changes in the pagerank random restart parameter affect the pagerank and relative ranking of web pages and whether these can be used as additional features for machine learning algorithms to increase their spam prediction rate. The following papers are related to what I am working on.

Web Content Categorization Using Link Information
Zoltan Gyongyi, Hector Garcia-Molina, and Jan Pedersen
http://infolab.stanford.edu/~zoltan/publications/gyongyi2006web.pdf

Making eigenvector-based reputation systems robust to boostings.
Technical report, USC, 2005
Hui Zhang, Ashish Goel, Ramesh Govindan, Kahn Mason, and Benjamin Van Roy.
ftp://ftp.usc.edu/pub/csinfo/tech-reports/papers/04-817.pdf

I am also reading parts of this doctoral dissertation

Detecting Colluders in PageRank - Finding Slow Mixing States in a Markov Chain
Kahn Mason. Doctoral Dissertation, Department of Management Science
and Engineering, Stanford University. September 2005.
http://www.stanford.edu/group/reputation/Mason_Thesis.pdf