

---

# VIRGIL PAVLU

vip@ccs.neu.edu  
www.ccs.neu.edu/home/vip  
617 921 5201

202 WVH Northeastern University  
360 Huntington Ave  
Boston MA 02115

---

INTERESTS Information Retrieval, Machine Learning, Applied Math, Artificial Intelligence

EMPLOYMENT Postdoctoral Associate, Northeastern University, August 2008 - August 2012. Participant on two awarded NSF funding proposals.  
Consultant on data mining problems, VOBA Solutions LLC, April 2009 - present

EDUCATION **PhD in Computer Science, Northeastern University, 2008.** Adviser: Javed Aslam.  
Thesis: *Evaluation for Large Scale Information Retrieval*; it was awarded the Northeastern University Completion Fellowship.  
**PhD Candidate in Computer Science, Dartmouth College 2001-2003**, adviser: Javed Aslam.  
Graduate level coursework: Algorithms, CS Theory, Systems, Machine Learning, Information Retrieval, AI, Programming Languages, Robotics, Computer Architecture, Information Theory, Parallel Computing, Functional Programming.  
**Bachelor in Mathematics (major) and Computer Science (minor)**, University of Bucharest, Romania. Graduation: June 2000.  
**Highschool:** National Romanian Math Team 1992-1996; National Math Olympiad first prize 1994-1995.

- RESEARCH
- **Score Distribution Models: Assumptions, Intuition, and Robustness to Score Manipulation** with Javed A. Aslam, Evangelos Kanoulas, Keshi Dai, SIGIR 2010
  - **Reusable Test Collections Through Experimental Design** with Ben Carterette, Evangelos Kanoulas, Hui Fang, SIGIR 2010
  - **Million Query Track 2009 Overview** with Ben Carterette, Evangelos Kanoulas, Hui Fang TREC 2009
  - **Variational Bayes for Modeling Score Distributions** with Javed A. Aslam, Evangelos Kanoulas, Keshi Dai, Information Retrieval Journal (Special Issue) 2009
  - **Modeling the Score Distributions of Relevant and Non-relevant Documents** with Javed A. Aslam, Evangelos Kanoulas, Keshi Dai, ICTIR 2009
  - **If I had a Million Queries** with James Allan, Ben Carterette, Javed A. Aslam, Evangelos Kanoulas, ECIR 2009
  - **Evaluation with Thousands of Queries** with James Allan, Ben Carterette, Javed A. Aslam, Evangelos Kanoulas, SIGIR 2008
  - **Million Query Track 2008 Overview** with James Allan, Ben Carterette, Javed A. Aslam, Evangelos Kanoulas, TREC 2008

- **Million Query Track 2007 Overview** with James Allan, Ben Carterette, Javed A. Aslam, Blagovest Dachev, Evangelos Kanoulas, TREC 2007
- **A Practical Sampling Strategy for Efficient Retrieval Evaluation**, with Javed Aslam, under review. [www.ccs.neu.edu/home/vip/study/publications\\_files/submitted.pdf](http://www.ccs.neu.edu/home/vip/study/publications_files/submitted.pdf)
- **Query Hardness Estimation Using Jensen-Shannon Divergence Among Multiple Scoring Functions**, with Javed Aslam, ECIR 2007.
- **The Hedge Algorithm for Metasearch at TREC 2006**, with Javed Aslam and Carlos Rei, TREC 2006.
- **Semi-supervised Data Organization for Interactive Anomaly Analysis**, with Javed Aslam and Sergey Bratus, ICMLA 2006.
- **A Statistical Method for System Evaluation Using Incomplete Judgments**, with Javed Aslam and Emine Yilmaz, SIGIR 2006.
- **The Maximum Entropy Method for Analyzing Retrieval Measures**, with Javed Aslam and Emine Yilmaz, SIGIR 2005.
- **A Geometric Interpretation of R-precision and Its Correlation with Average Precision**, with Javed Aslam and Emine Yilmaz, SIGIR 2005.
- **Measure-based Metasearch**, with Javed Aslam and Emine Yilmaz, SIGIR 2005.
- **A Sampling Technique for Efficiently Estimating Measures of Query Retrieval Performance Using Incomplete Judgments**, with Javed Aslam and Emine Yilmaz, ICML 2005 Workshop on Learning with Partially Classified Training Data.
- **A Unified Model for Metasearch, Pooling, and System Evaluation**, with Javed Aslam and Robert Savell, CIKM 2003.
- **A Unified Model for Metasearch and the Efficient Evaluation of Retrieval Systems via the Hedge Algorithm**, with Javed Aslam and Robert Savell, SIGIR 2003.

ACTIVITY Workshop organizer, SIGIR 2010  
 Local Arrangements Chair, SIGIR 2009  
 TREC coordinator for Million Query Track, 2009  
 Development of Evaluation method used by Million Query Track, TREC 2007-2008  
 Program Committee SIGIR 2007,2008,2010,2011; ECIR 2011  
 External Reviewer SIGIR 2005-2006, CIKM 2005-2007  
 ACM member

TEACHING **Instructor of Record**, Northeastern Univ  
 Information Retrieval (undergraduate) 2005-2008  
 Information Retrieval (graduate) 2007,2009  
 Machine Learning (graduate) 2008,2010  
 Algorithms (graduate) 2009,2010  
[www.ccs.neu.edu/home/vip/teach/](http://www.ccs.neu.edu/home/vip/teach/)  
**Teaching assistant**  
 Dartmouth College: Algorithms (2001,2002), Theory (2002), Discrete Math (2001,2003), Computer Architecture (2002)

Northeastern Univ: Theory (2003), Information Retrieval (2005)  
University of Bucharest, Romania: Lab assistant, TA for discrete structures class.

#### SKILLS

**outstanding:** teaching, presentation software

**very good:** applied math, design, lecturing, problem-solving, machine learning, data analysis, optimization, Matlab, indexing, retrieval & matching, MS Office, Mac Office

**good:** algorithm design, compression schemes, text processing, hacks & improvisations, C/C++, Perl, MS Visual Studio, Latex, Windows OS, MAC OS, Linux OS, repair systems (hardware and software), interfaces

**familiar with:** audio/video formats, Java, html, shell scripting, image processing, photography

HOBBIES Geometry, Bridge, Music and Percussion, Wood working, Table Tennis, Strategy Games

Last Updated : December 13, 2010