

# Hilmi Yildirim

AE 106  
110 8th St.  
Troy, NY 12180

518-605-8872  
[yildih@rpi.edu](mailto:yildih@rpi.edu)  
<http://www.cs.rpi.edu/~yildih2/>

## Education

- **PhD, Computer Science** January 2008 - August 2011 (*Expected*)  
*Rensselaer Polytechnic Institute* Troy, NY
  - Advisor : Mohammed Javeed Zaki
  - Thesis Title : Graph Indexing for Reachability Queries
  - cGPA : 3.83/4.00
- **M.S, Computer Science** August 2005 - December 2007  
*Rensselaer Polytechnic Institute* Troy, NY
  - Thesis Title: A Random Walk Method For Handling Sparse Training Set in Collaborative Filtering
  - Advisor: Mukkai S. Krishnamoorthy
  - cGPA : 3.75/4.00
- **B.S, Computer Engineering** 2001-2005  
*Bilkent University* Ankara, Turkey
  - Graduated with High Honors, cGPA : 3.54 / 4.00, and Dean's Honour List each year.
  - Senior Project: Graph Algorithms for Advanced Querying in PATIKApr

## Industrial Experience

- **Internship** May 2010 - August 2010  
*Group: Enterprise Manager - Application Performance Management* Oracle Inc., Redwood City, CA
  - JVM Heap Diagnostics: Designed and developed scalable methods for visualization and analysis of memory leaks
- **Internship in Data Mining** May 2008 - August 2008  
*Group: Advanced Concepts & Technologies* Pitney Bowes Inc., Shelton, CT
  - Developed practical techniques for *sentiment analysis* from unstructured text data, classified customers according to their satisfaction level from auto-decrypted phone conversations with the customer service
  - Developed a web application (database front end) for the database of customer surveys with advanced querying facility. Application is developed in Java connecting to an Oracle server.

## Academic Experience

- **Research Assistant** Fall 2008 - August 2011 (*Expected*)  
*Data Mining Lab* Rensselaer Polytechnic Institute, Troy, NY
  - Developed novel scalable techniques for indexing directed graphs for reachability queries
    - \* Reachability Query : Is node  $x$  is reachable from node  $y$  in a directed graph?
    - \* Emerging real world applications, as diverse as XML databases, GIS, web mining, social network analysis, ontologies, and bioinformatics.
    - \* Applied our algorithm to large scale graphs of million node size and it outperformed most of the existing methods.
  - Developed a novel item-oriented collaborative filtering algorithm for sparse datasets
    - \* Item-oriented methods discover item-item relationships from the training data and use these relations to compute predictions. When the training set is not dense enough, these systems suffer from producing high quality predictions. I developed a random walk method in which each user is modelled as a random-buyer who takes finite-length random walks in the item graph. With this method, the system becomes able to capture hidden relations without increasing the order of computational complexity. Published in Recsys 2008.

- Developed a novel scalable technique for capturing arbitrary shaped clusters in high dimensional data
  - \* Spatial clustering has been applied to data from astronomy, meteorology, epidemiology, seismology, etc.
  - \* Existing methods require careful parameter tuning and are not scalable.
  - \* Proposed a scalable and robust method for spatial clustering that is based on the hypothesis that a spatial cluster originates from a set of core points within the cluster that form the backbone or the intrinsic shape of the cluster.
- Parallelized “Genome Scale Indexing: Disk-based Suffix Trees - TRELLIS” in Blue Gene Supercomputer
  - \* Trellis is a disk-based suffix tree indexing method (with suffix links) that is capable of indexing the entire human genome on a commodity PC with limited memory. Implemented parallelized version of Trellis using MPI on RPI’s Blue Gene Supercomputer which has 32768 processors.

• **Research Assistant** June 2006 - December 2007  
*Culturally-Situated Design Tools* **Rensselaer Polytechnic Institute, Troy, NY**

- Implemented a Fractal Generator Tool applet in JAVA
- Designed a programming language and a development environment for the *VirtualBreakdancer* tool
- Designed and implemented a 3D Virtual Breakdancer applet written in JAVA with OpenGL available at <http://www.ccd.rpi.edu/Eglash/csdt/subcult/brdance/index.html>

• **Teaching Assistant** August 2005 - December 2008  
**Rensselaer Polytechnic Institute, Troy, NY**

- Taught *Programming Languages, Models of Computation, Database Systems and Operating Systems*

• **Undergraduate Researcher** May 2004 - Aug. 2004  
*PATIKA @ Bilkent Center for Bioinformatics* **Bilkent, Ankara, Turkey**

- Designed and implemented graph algorithms for advanced querying in a pathway database

• **Tutor and Lab Assistant** August 2004 - May 2005  
**Bilkent, Ankara, Turkey**

- Tutored *Introduction to Programming and Algorithms with Matlab and JAVA*

## Awards

Full Scholarship awarded by TUBITAK (The Scientific and Technical Research Council of Turkey)	2001-2005
Full Scholarship awarded by Bilkent University	2001-2005
Ranked 255th in the University Entrance Examination among 1.5 million students	2001
Participated in the 13th International Olympiads in Informatics (IOI'01) in <i>Tampere, Finland</i>	2001
Silver medal at the 9th Balkan Olympiads in Informatics (BOI'01) in <i>Durres, Albania</i>	2001
Silver medal at the National Olympiads in Informatics organized by TUBITAK	2000
Ranked 410th in the National Science Schools Entrance Examination among 750000 students	1998

## Publications

1. Hilmi Yildirim, Vineet Chaoji, and Mohammed Zaki “*GRAIL: A Scalable Index for Reachability Queries in Very Large Graphs*”, (submitted to VLDB Journal)
2. Vineet Chaoji, Li Geng, Hilmi Yildirim, and Mohammed Zaki “*ABACUS: Mining Arbitrary Shaped Clusters from Large Datasets based on Backbone Identification*”, SIAM Data Mining (SDM), Mesa, Arizona, April 2011
3. M. A. Hasan, H. Yildirim, A. Chakrabarty: “*SONNET: Efficient Approximate Nearest Neighbor using Multi-Core*”, IEEE International Conference on Data Mining (ICDM), Sydney, Australia, December 2010
4. Hilmi Yildirim, Vineet Chaoji, and Mohammed Zaki “*GRAIL: Scalable Reachability Index for Large Graphs*”, 36th International Conference on Very Large Data Bases (VLDB), Singapore, September 2010
5. Hilmi Yildirim “*Graph Indexing for Reachability Queries*”, 26th IEEE International Conference on Data Engineering Phd Workshop, Long Beach, California, March 6, 2010
6. Hilmi Yildirim and Mukkai S. Krishnamoorthy “*A Random Walk Method for Alleviating the Sparsity Problem in Collaborative Filtering*”, 2nd ACM International Conference on Recommender Systems, Lausanne, Switzerland, October 23-25, 2008
7. Hilmi Yildirim, Mukkai S. Krishnamoorthy, Narsingh Deo “*A Study of the Sudoku Graph Family*” Southeastern International Conference on Combinatorics, Graph Theory, and Computing, Boca Raton, USA, March, 2008

8. MS Thesis: “A Random Walk Method For Handling Sparse Training Set In Collaborative Filtering”, Department of Computer Science, Rensselaer Polytechnic Institute, December 2007
9. C. Aksay, F. Arik, E. Ataer, H. Yildirim et al. “PATIKAwEB: A Web Service for Querying, Visualizing, and Analyzing a Graph Based Pathway Database”, ISMB 2005, Detroit, USA, June 26 - 29, 2005
10. C. Aksay, A. Ayaz, O. Babur, H. Yildirim et al. “PATIKA: An informatics infrastructure for cellular networks”, ISMB 2004, Glasgow, Scotland, July 31 - August 4, 2004.

## Course Projects

- *Parallelized Suffix Tree Construction* **High Performance Parallel & Distributed Computing**  
Parallelization of a very recent genome-scale indexing algorithm using MPI on RPI's supercomputer which has 32768 processors. Experiments are done on 2048 processors.
- *HMM Analysis of a Genome* **Algorithms in Bioinformatics**  
Implementation of Hidden Markov Model methods for exploring meaningful regions of the genome of *Pyrococcus Horikoshii*
- *Operating System Simulation* **Computer Operating Systems**  
Implementation of a multi-processor, multi-disk operating system simulator in JAVA, and analysis of the affects of parameters(#CPUS, #disks, cache size,etc.) on the performance of the system
- *Turn Based Strategy Game* **Object-Oriented Software Engineering**  
Design and implementation of a fully object-oriented turn-based strategy game written in JAVA, my focus was on client-server interaction
- *Animated Spider* **Computer Graphics**  
Implemented in JAVA using *GL4JAVA* library
- *Online Library Administration and Circulation System* **Database Systems**  
A 3-person team project written in C#, ASP and MSSQL
- *Network Emulator, Heap Management Library, Floppy File System* **Operating Systems**

## Skills

**Languages:** Java, C/C++, L<sup>A</sup>T<sub>E</sub>X, Matlab, Octave, Perl, Python, SQL, PostgreSQL, MPI, OpenGL

**Operating Systems:** Unix/Linux, Solaris, Windows, Mac OS

**Applications:** Eclipse, Netbeans, JIdea, Oracle, MS Office, OpenOffice, JBuilder, CVS, Subversion

## Extracurricular Activities

**Turkish Cultural Center Albany:** Founder and Treasurer of Turkish Cultural Center Albany

Responsible for bookkeeping using Quickbooks, maintaining website, prepare financial statements and projections, oversee tax documents, help organizing cultural and educational activities.

## References

Prof. Mohammed J. Zaki  
zaki@cs.rpi.edu

Assoc. Prof. Mukkai S. Krishnamoorthy  
moorthy@cs.rpi.edu

Asst. Prof. Elliot Anshelevich  
eanshel@cs.rpi.edu