MOHAMED A. SOLIMAN

David R. Cheriton School of Computer Science
University of Waterloo
200 University Ave. West
Waterloo, Ontario, N2L 3G1, Canada

m2ali@cs.uwaterloo.ca
http://www.cs.uwaterloo.ca/~m2ali
Phone: +1 519 888 4567 Ext. 33143

Research Interests

Database Systems, ranking and top-k retrieval, probabilistic databases, data cleaning, preference management, information extraction.

Education

- **David R. Cheriton School of Computer Science, University of Waterloo**, **Ontario, Canada**
  Ph.D. student, started 2005
  Advisor: Ihab Ilyas
- **M.S. in Computer Science, Alexandria University, Egypt, 2004**
  Thesis: *Web Documents Classification Using Text, Anchor, Title, and Metadata Information*
- **B.S. in Computer Science, Alexandria University, Egypt, 2000**
  First degree of honor

Honors and Awards

- Ontario Graduate Scholarship (OGS), Ontario Government, 2009-2010
- Cheriton Excellence Scholarship, University of Waterloo, 2007-2009
- Research in Motion (RIM) Graduate Scholarship, University of Waterloo, 2010
- Graduate Merit Scholarship, University of Waterloo, 2006
- Entrance Scholarship, University of Waterloo, 2005
- B.S. first degree of honor, Department of Computer Science, Alexandria University, Egypt, 2000

Systems

- **URank:** Ranking and Aggregation in Uncertain Database Systems.

http://www.cs.uwaterloo.ca/~ilyas/URank/

Lead the design and implementation of the whole system, and contributed in building new models to integrate uncertainty and ranking awareness in a full-fledged database system.
System demonstrations are in SIGMOD'07 and ICDE'10.

- **RankDB:** Rank-aware Query Processing in Database Systems.

http://www.cs.uwaterloo.ca/~ilyas/RankDB/

Extended DBMS plan enumeration algorithm to consider ranking as a first class citizen. Extended join operators to efficiently handle ranking requirements of SQL queries.
System demonstration is in VLDB'05.

Research and Professional Experience

University of Waterloo, Waterloo, ON, Canada ***2005 - Present***

Graduate Research Assistant, School of Computer Science

Formulated new definitions for ranking and ranking-aggregate queries in probabilistic databases. Introduced new ranking models linking probabilistic data with partial orders. Designed and implemented a new Web mashup system that supports ranking of uncertain data. Developed new query processing algorithms for computing probabilistic ranking queries. Implemented uncertainty and rank-aware query operators inside PostgreSQL. Designed a model and query processing algorithms for Pareto-optimal Web interactions based on Game Theory. Implemented a new rank-aware query plan enumeration algorithm inside PostgreSQL.

Yahoo! Labs, Santa Clara, CA, USA ***Summer 2009***

Intern at the Community Systems Group

Mentor: *Nilesh Dalvi,* Manager**:** *Philip Bohannon*

Was an active member in the Purple Socially Oriented Extraction (PSOX) project, which focuses on large scale information extraction and entity matching on the Web. Carried out several research tasks including designing new techniques for wrapper induction from noisy examples, using entity matching to detect errors and debug information extraction techniques, using Machine Learning concepts to model and query information extraction rules under noise.

IBM Almaden Research Center, CA, USA ***Fall 2006***

Intern at the Advanced Database Solutions Department

Mentor: *Volker Markl,* Manager**:** *Guy Lohman*

Was an active member in the DAMIA project which focuses on building mashups for Intranet applications. Carried out several research tasks including modeling uncertainty in mashup construction, integrating text annotators of Web documents in mashup data flows, and implementing probabilistic mashup operators.

Alexandria University, Department of Computer Science, Alexandria, Egypt ***2001-2004***

Graduate Research Assistant

Designed a new classification algorithm for Web documents based on association rules mined from text, title, anchor, and metadata information.

Information Systems Technology, Alexandria, Egypt ***2000-2001***

Developer and System Analyst

Worked as a part time system developer at "Information System Technology" (IST), Egypt. Participated in a large database project that manages hotel business using MS SQL Server. Work assignments included data modeling, system-level implementation, testing and debugging, and customer interaction.

Teaching Experience

University of Waterloo, Waterloo, ON, Canada ***2005 - Present***

Teaching Assistant

TA for different graduate/undergraduate level courses (CS448/648, CS348, CS125) in the areas of databases and programming principles. Shared responsibility for tutorials, exams, assignments, and marking. Duties at various

times have included giving lecturers, interaction with students, marking, news groups monitoring/maintenance, assignment submission control.

Alexandria University, Department of Computer Science, Alexandria, Egypt ***2001***
Teaching Assistant

TA for several Computer Science courses including Internet Programming, Operating Systems, and Computer Applications. Worked as an instructor for computer training courses including "SQL Server administration", "Visual Basic Programming", "Microsoft.NET Framework", "Programming with Java", "Office and Windows Applications".

Publications

Refereed Journal Publications

1. Mohamed A. Soliman, Ihab F. Ilyas, and Shalev Ben-David, "Supporting Ranking Queries on Uncertain and Incomplete Data", *to appear in the VLDB Journal*, 2010.

2. Mohamed A. Soliman, Ihab F. Ilyas, Kevin C. Chang, "Probabilistic Top-k and Ranking-Aggregate Queries*", in the ACM Transactions on Database Systems,* ***TODS***, Volume 33, Issue 3, August 2008, pp. 1-54.

3. Ihab F. Ilyas, George Beskales, Mohamed A. Soliman, "A Survey of Top-k Query Processing Techniques in Relational Database Systems", *in* ***the ACM Computing Surveys****, Volume 40, Issue 4, October 2008, Article 11*.

Refereed Conference Papers

4. Mohamed A. Soliman, Ihab F. Ilyas, and Mina Saleeb, "Building Ranked Mashups of Unstructured Sources with Uncertain Information", *to appear in proceedings of the 36th International Conference on Very Large Data Bases,* ***VLDB 2010***, Singapore.

5. George Beskales, Mohamed A. Soliman, Ihab F. Ilyas, and Shai Ben-David, "Modeling and Querying Possible Repairs in Duplicate Detection", *in Proceedings of the 35th International Conference on Very Large Data Bases,* ***PVLDB*** 2(1), Lyon, France, pp. 598-609.

6. Mohamed A. Soliman, Ihab F. Ilyas, "Ranking with Uncertain Scores", *in Proceedings of the 25th IEEE International Conference on Data Engineering*, ***ICDE 2009***, Shanghai, China, pp. 317-328.

7. George Beskales, Mohamed A. Soliman, Ihab F. Ilyas, "Efficient Search for the Top-k Probable Nearest Neighbors in Uncertain Databases", *in Proceedings of the 34th International Conference on Very Large Data Bases,* ***PVLDB*** 1(1), Auckland, New Zealand, pp. 326-339.

8. Mohamed A. Soliman, Ihab F. Ilyas, Kevin C. Chang, "Top-k Query Processing in Uncertain Databases", *in Proceedings of the 23rd IEEE International Conference on Data Engineering,* ***ICDE 2007****,* Istanbul, Turkey, pp. 896-905.

9. Mohamed A. Soliman, Ihab F. Ilyas, Nick Koudas, "Finding Skyline and Top-k Bargaining Solutions", *(poster paper)*, *in Proceedings of the 23rd IEEE International Conference on Data Engineering,* ***ICDE 2007****,* Istanbul, Turkey, pp. 1263-1267.

10. Mohamed Fathi, Noha Adly, Magdy Nagi, "Web Documents Classification Using Text, Anchor, Title and Metadata information", Master Thesis, Computer Science Department, Alexandria University, Egypt, 2004, *in Proceedings of the International Conference on Computer Science, Software Engineering, Information Technology, e-Business and Applications*, ***CSITeA 2004*****,** Cairo, Egypt, pp. 445-452.

Refereed System Demonstrations

11. Mohamed A. Soliman, Mina Saleeb, and Ihab F. Ilyas, "MashRank: Towards Uncertainty-Aware and Rank-Aware Mashups", *to appear in Proceedings of the 26th IEEE International Conference on Data Engineering*, ***ICDE 2010***, Long Beach, California, USA.

12. George Beskales, Mohamed A. Soliman, Ihab F. Ilyas, and Shai Ben-David, "ProbClean: A Probabilistic Duplicate Detection System", *to appear in Proceedings of the 26th IEEE International Conference on Data Engineering*, ***ICDE 2010***, Long Beach, California, USA.

13. Mohamed A. Soliman, Ihab F. Ilyas, Kevin C. Chang, "URank: Formulation and Efficient Evaluation of Top-k Queries in Uncertain Databases", *in Proceedings of the* ***2007 ACM SIGMOD*** *Conference on Management of Data*, Beijing, China, pp. 1082-1084.

14. Chengkai Li, Mohamed A. Soliman, Kevin Chen-Chuan Chang, Ihab F. Ilyas, "RankSQL: Supporting Ranking Queries in Relational Database Management Systems", *in Proceedings of the 31st International Conference on Very Large Data Bases,* ***VLDB 2005***, Trondheim, Norway, pp. 1342-1345.

Computer Skills

- Programming Languages: C/C++, Java, Pascal, VB, C#, Client/Server Side Scripting (VB/Java script, JSP, ASP, PHP), some knowledge of Unix scripts.
- Operating Systems: Unix/Linux, Windows.
- Excellent knowledge of the Microsoft .NET platform, and MS SQL Server.
- Good knowledge of the internals of PostgreSQL on the level of system core implementation.

Other Information

- Please visit http://www.cs.uwaterloo.ca/~m2ali for more information and copies of representative publications.
- Born February 12, 1978 in Alexandria, Egypt. Citizen of Egypt. Permanent Resident of Canada.