

## Tamer Mohamed Elsayed

### EDUCATION

---

**Ph.D. in Computer Science, August 2009**

*University of Maryland-College Park, MD, US*  
Main Research Area: Information Retrieval  
Dissertation Topic: Identity Resolution in Email Collections  
Advisor: Prof. Douglas W. Oard  
Award: Graduate Fellowship, Fall 2002 - Spring 2004  
GPA: 3.8

**M.Sc. in Computer Science, May 2005**

*University of Maryland-College Park, MD, US*  
GPA: 3.8

**M.Sc. in Computer Science, July 2001**

*Alexandria University, Egypt*  
Thesis Topic: Feedback Consolidation in Point-to-Multipoint Connections of ABR Service in ATM Networks.  
Advisor: Prof. Nazih El-Derini  
GPA: 4.0

**B.Sc. in Computer Science and Automatic Control, June 1997**

*Alexandria University, Egypt*  
Graduation Grade: Excellent with degree of honor  
Awards: Faculty Certificate of Honor 1993-1997.  
Graduation Project: Human Face Recognition.  
GPA: 3.7

### EMPLOYMENT

---

**Assistant Professor, September 2012 – Present**

*Computer Science and Engineering Dept., College of Engineering  
Qatar University, Doha, Qatar*

**Researcher, September 2011 – August 2012**

*Microsoft Advanced Technology Lab  
Cairo, Egypt*

**Post-Doctoral Fellow, September 2010 – August 2011**

*Mathematical and Computer Sciences and Engineering Division  
King Abdullah University of Science and Technology (KAUST), Saudi Arabia*

**Post-Doctoral Researcher, September 2009 – August 2010**

*iSchool, Cloud Computing Center  
University of Maryland-College Park, MD, US*

**[Summer Internship] Software Engineer Intern, June 2007 – August 2007**

*Google, Mountain View, CA, US*

**Graduate Research Assistant, June 2004 – August 2009**

*Computer Science Department  
University of Maryland-College Park, MD, US*

**Graduate Teaching Assistant, August 2002 – May 2004**

*Computer Science Department  
University of Maryland-College Park, MD, US*

**Graduate Teaching Assistant, January 1998 – July 2002**

*Department of Computers and Automatic Control  
Alexandria University, Alexandria, Egypt*

**Software Engineer, August 1997 – July 2002**

*KME Computools, Alexandria, Egypt*

**AWARDS & ACHIEVEMENTS**

---

**Best Paper Award at AIRS 2015**

Got best paper award at Asia Information Retrieval Societies Conference (AIRS) 2015 for the paper titled “Improving Tweet Timeline Generation by Predicting Optimal Retrieval Depth” co-authored with my prospective PhD student Maram Hasanain and Walid Magdy (QCRI).

**Ranked 2nd at TREC Microblog Track 2015**

Our QU team (Reem Suwaileh, Maram Hasanain, and me) got ranked 2<sup>nd</sup> at TREC 2015, Microblog track, realtime filtering task, push notification scenario among 14 participating international teams.

**Best CSE Senior Project Award (2014/2015)**

My senior students Abdelrazek Ibrahim, Karim Fahmi, and Ahemd Abdelhamid team won the “Best CSE Senior Project Award” of academic year 2014/2015 with their senior (graduation) project titled “Smart Shopping Cart” that I supervised.

**First MSc with Distinction in QU-CENG History (Maram Hasanain)**

My student Maram Hasanain has successfully defended her MSc thesis titled “Query Performance Prediction for Microblog Search” and got her degree \*with Distinction\* (with consensus of all members of the examination committee) in the first time in the history of College of Engineering at Qatar University! Examination committee included Prof. Ali Jaoua (QU), Dr. Abdelkarim Erradi (QU), and Prof. Franciska de Jong (University of Twente).

**RESEARCH INTERESTS**

---

**INFORMATION RETRIEVAL & BIG DATA**

- Designing and evaluation of search engines/information retrieval systems
- Big data analytics with emphasis on text and web scale
- Searching informal media (e.g., emails, tweets, blogs)

**PUBLICATIONS**

---

*JOURNAL PUBLICATIONS*

**Unsupervised Adaptive Microblog Filtering for Broad Dynamic Topics**

*Walid Magdy and Tamer Elsayed*

Information Processing and Management. Elsevier. Vol. 52(4), pp. 513-528, July 2016.

**CloudFlow: Data-aware Programming Model for Cloud Workflow Applications on Modern HPC Systems**

*Fan Zhang, Qutaibah. M. Malluhi, Tamer Elsayed, Samee U. Khan, Keqin Li*

Future Generation Computer Systems Journal, Vol. 51, pp. 98–110, Elsevier, October 2015.

**Making Sense of Archived Email: Exploring the Enron Collection with NetLense**

*Hyunmo Kang, Catherine Plaisant, Tamer Elsayed, and Douglas W. Oard*

Journal of the American Society for Information Science and Technology (JASIST), 61: 723–744, April 2010.

**IBN: A Communication Paradigm for Mobile Applications**

*Moustafa Youssef, Tamer Elsayed, Mohamed Hussein, Tamer Nadeem, Adel Youssef, and Liviu Iftode*

Mobile Computing and Communications Review. Vol. 7, No. 4, pp. 66-67, October 2003.

## CONFERENCE PUBLICATIONS

### **EveTAR: A New Test Collection for Event Detection in Arabic Tweets**

*Hind Almerekhi, Maram Hasanain, and Tamer Elsayed*

Proceedings of the 39<sup>th</sup> annual international ACM SIGIR conference on Research and development in information retrieval: SIGIR '16, Pisa, Italy, July 2016 (to appear).

### **ArabicWeb16: A New Crawl for Today's Arabic Web**

*Reem Suwaileh, Mucahid Kultlu, Nihal Fathima, Tamer Elsayed, and Matthew Lease*

Proceedings of the 39<sup>th</sup> annual international ACM SIGIR conference on Research and development in information retrieval: SIGIR '16, Pisa, Italy, July 2016 (to appear).

### **On the Evaluation of Tweet Timeline Generation Task**

*Walid Magdy, Maram Hasanain, and Tamer Elsayed*

Proceedings of the 38<sup>th</sup> European Conference on Information Retrieval: ECIR 2016, pp. 648-653, Padua, Italy, March 2016.

### **Detecting Automatically-Generated Arabic Tweets**

*Hind Almerekhi and Tamer Elsayed*

Proceedings of the 11<sup>th</sup> Asia Information Retrieval Societies Conference: AIRs 2015, pp. 123-134, Brisbane, Queensland, Australia, Springer, December 2015.

### **Improving Tweet Timeline Generation by Predicting Optimal Retrieval Depth**

*Maram Hasanain, Tamer Elsayed, and Walid Magdy*

Proceedings of the 11<sup>th</sup> Asia Information Retrieval Societies Conference: AIRs 2015, pp. 135-146, Brisbane, Queensland, Australia, Springer, December 2015.

### **Identification of Answer-Seeking Questions in Arabic Microblogs**

*Maram Hasanain, Tamer Elsayed, and Walid Magdy*

Proceedings of the 23<sup>rd</sup> ACM Conference on Information and Knowledge Management: CIKM'14, pp. 1839-1842, November 2014.

### **Adaptive Method for Following Dynamic Topics on Twitter**

*Walid Magdy and Tamer Elsayed*

Proceedings of the 8<sup>th</sup> International Conference on Weblogs and Social Media (ICWSM), pp. 335-345, June 2014.

### **iHadoop: Asynchronous Iterations for MapReduce**

*Eslam Elnikety, Tamer Elsayed, and Hany Ramadan*

Proceedings of the 3<sup>rd</sup> IEEE annual international conference Cloud Computing Technology and Science: CloudCom '11, pp. 81-90, November 2011.

### **When Close Enough Is Good Enough: Approximate Positional Indexes for Efficient Ranked Retrieval**

*Tamer Elsayed, Jimmy Lin, and Don Metzler*

Proceedings of the 20<sup>th</sup> ACM Conference on Information and Knowledge Management: CIKM'11, pp. 1993-1996, October 2011.

### **No Free Lunch: Brute-force vs. Locality-Sensitive Hashing for Cross-lingual Pairwise Similarity**

*Ferhan Ture, Tamer Elsayed, and Jimmy Lin*

Proceedings of the 34<sup>th</sup> annual international ACM SIGIR conference on Research and development in information retrieval: SIGIR '11, pp. 943-952, July 2011.

### **Pseudo Test Collections for Learning Web Search Ranking Functions**

*Nima Asadi, Don Metzler, Tamer Elsayed, and Jimmy Lin*

Proceedings of the 34<sup>th</sup> annual international ACM SIGIR conference on Research and development in information retrieval: SIGIR '11, pp. 1073-1082, July 2011.

### **Predicting Author Blog Channels with High Value Future Posts for Monitoring**

*Shanchan Wu, Tamer Elsayed, William Rand, and Louiqa Raschid*

Proceedings of the 25<sup>th</sup> annual conference on AI: AAI '11, pp. 1261-1266, August 2011.

### **Arabic Cross-Document Coreference Detection**

*Asad Sayeed, Tamer Elsayed, Nikesh Garera, David Alexander, Tan Xu, Douglas W. Oard, David Yarowsky, and Christine Piatko*

Proceedings of the 47<sup>th</sup> Annual meeting of the Association of Computational Linguistics (ACL), pp. 357-360, August 2009.

**Resolving Personal Names in Email Using Context Expansion**

*Tamer Elsayed, Douglas W. Oard, and Galileo Namata*

Proceedings of the 46<sup>th</sup> Annual meeting of the Association of Computational Linguistics: Human Language Technologies ACL/HLT-08, pp. 941-949, June 2008.

**Pairwise Document Similarity in Large Collections with MapReduce**

*Tamer Elsayed, Jimmy Lin, and Douglas W. Oard*

Proceedings of the 46<sup>th</sup> Annual meeting of the Association of Computational Linguistics: Human Language Technologies ACL/HLT-08: Short Papers, pp. 265-268, June 2008.

**Knowledge Representation from Information Extraction**

*Tan Xue, Douglas W. Oard, Tamer Elsayed, and Asad Sayeed*

Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries: JCDL '08, pp. 475-475, June 2008.

**Modeling Identity in Archival Collections of Email: A Preliminary Study**

*Tamer Elsayed and Douglas W. Oard*

Proceedings of the 3<sup>rd</sup> Conference on Email and Anti-Spam (CEAS), pp. 95-103, July 2006.

**On Evaluation of Adaptive Topic Tracking Systems**

*Tamer Elsayed and Douglas W. Oard*

Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval: SIGIR '05, pp. 597-598, August 2005.

**ATP: Autonomous Transport Protocol**

*Tamer Elsayed, Mohamed Hussein, Moustafa Youssef, Tamer Nadeem, Adel Youssef and Liviu Iftode*

4-page Paper at IEEE Midwest Symposium on Circuits and Systems, Cairo, Egypt, December 2003.

**Autonomous Transport Protocol**

*Tamer Elsayed, Mohamed Hussein, Moustafa Youssef, Tamer Nadeem, Adel Youssef and Liviu Iftode*

Research Poster/1-page Extended Abstract at the Eleventh Annual International Conference on Network Protocols, (ICNP) 2003, Atlanta, GA, USA, November 2003.

**IBN: A Communication Paradigm for Mobile Applications**

*Moustafa Youssef, Tamer Elsayed, Mohamed Hussein, Tamer Nadeem, Adel Youssef and Liviu Iftode*

Research Poster/2-page Short Paper at the Ninth Annual International Conference on Mobile Computing and Networking, (MobiCom) 2003, San Diego, CA, USA, September 2003.

**Improved Algorithm for Feedback Consolidation in Multicast ABR Connections**

*Tamer Elsayed, Mohamed Elderini, Mohamed Selim, and Magdi Ahmed*

4-page Paper at the 1st IEEE International Symposium on Signal Processing and Information Technology, Cairo, Egypt, December 2001.

**PHD DISSERTATION**

**Identity Resolution in Email Collections**

Computer Science Dept., University of Maryland, College Park, MD, US, August 2009.

**MASTER THESIS**

**Feedback Consolidation in Point-to-Multipoint Connections of ABR Service in ATM Networks**

Department of Computers and Automatic Control, University of Alexandria, Alexandria, Egypt, June 2001.

#### **REFEREED WORKSHOP PUBLICATIONS**

##### **QU at SemEval 2016 Task 3: Learning to Rank on Arabic Question Answering Forums with Word Embedding**

*Rana Malhas, Marwan Torki, and Tamer Elsayed*

Proceedings of the 10<sup>th</sup> International Workshop on Semantic Evaluation, SemEval '16, San Diego, California, June 2016 (to appear).

##### **What Questions Do Journalists Ask on Twitter?**

*Maram Hasanain, Mossaab Bagdouri, Tamer Elsayed, and Douglas W. Oard*

Proceedings of the Social Media in the Newsroom Workshop (SMNews'16) at International Conference on Web and Social Media (ICWSM). pp. 127-134. Cologne, Germany, May 2016.

##### **Query Performance Prediction for Microblog Search: A Preliminary Study**

*Maram Hasanain, Rana Malhas, and Tamer Elsayed*

Proceedings of the SIGIR International Workshop on Social Media Retrieval and Analysis (SoMeRA'14), pp. 1-6, July 2014.

##### **ConMR: Concurrent MapReduce Programming Model for Large Scale Shared-Data Applications**

*Fan Zhang, Qutaibah Malluhi, and Tamer Elsayed*

Sixth International Workshop on Parallel Programming Models and Systems Software for High-End Computing (P2S2), 2013.

##### **Cross-Document Coreference Resolution: A Key Technology for Learning by Reading**

*James Mayfield, David Alexander, Bonnie Dorr, Jason Eisner, Tamer Elsayed, Tim Finin, Clay Fink, Marjorie Freedman, Nikesh Garera, Paul McNamee, Saif Mohammad, Douglas Oard, Christine Piatko, Asad Sayeed, Zareen Syed, Ralph Weischedel, Tan Xu and David Yarowsky.*

AAAI 2009 Spring Symposium, Workshop of Learning by Reading and Learning to Read.

#### **EVALUATION CONFERENCE PAPERS<sup>1</sup>**

##### **QU at TREC-2015: Building Real-Time Systems for Tweet Filtering and Question Answering**

*Reem Suwaileh, Maram Hasanain, Marwan Torki, and Tamer Elsayed*

Proceedings of the 24<sup>th</sup> Text REtrieval Conference (TREC 2015), Gaithersburg, Maryland, USA, NIST, January 2016.

##### **QU at TREC-2014: Online Clustering with Temporal and Topical Expansion for Tweet Timeline Generation**

*Maram Hasanain and Tamer Elsayed*

Proceedings of the 23<sup>rd</sup> Text REtrieval Conference (TREC 2014), Gaithersburg, Maryland, USA, NIST, January 2015.

##### **QU at TREC-2013: Expansion Experiments for Microblog Ad-hoc Search**

*Maram Hasanain, Latifa Al-Marri, and Tamer Elsayed*

Proceedings of the 22<sup>nd</sup> Text REtrieval Conference (TREC 2013), Gaithersburg, Maryland, USA, NIST, November 2013.

##### **BEST of KAUST at TREC-2011: Building Effective Search in Twitter**

*Jinling Jiang, Lailatul Hidayah, Tamer Elsayed, and Hany Ramadan*

Proceedings of the 20<sup>th</sup> Text REtrieval Conference (TREC 2011), Gaithersburg, Maryland, NIST, November 2011.

##### **UMD and USC/ISI: TREC 2010 Web Track Experiments with Ivory**

*Tamer Elsayed, Nima Asadi, Donald Metzler, Lidan Wang, and Jimmy Lin*

Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010), Gaithersburg, Maryland, NIST, November 2010.

---

<sup>1</sup> Published but not refereed.

**Of Ivory and Smurfs: Loxodontan MapReduce Experiments for Web Search**

*Jimmy Lin, Donald Metzler, Tamer Elsayed, and Lidan Wang*

Proceedings of the Eighteenth Text REtrieval Conference (TREC 2009), Gaithersburg, Maryland, NIST, November 2009.

**TREC 2006 at Maryland: Blog, Enterprise, Legal and QA Tracks**

*Douglas W. Oard, Tamer Elsayed, Jianqiang Wang, Yejun Wu, Pengyi Zhang, Eileen Abels, Jimmy Lin, and Dagbert Soergel*

Proceedings of the Fifteenth Text REtrieval Conference (TREC 2006), pp. 199-214, Gaithersburg, Maryland, NIST, November 2006.

**TDT-2004: Adaptive Topic Tracking At Maryland**

*Tamer Elsayed, Douglas W. Oard, David Doermann, and Gary Kuhn*

In Working Notes of Topic Detection and Tracking (TDT) Evaluation, Gaithersburg, Maryland, NIST, November 2004.

**TECHNICAL REPORTS**

**Brute-Force Approaches to Batch Retrieval: Scalable Indexing with MapReduce, or Why Bother?**

*Tamer Elsayed, Ferhan Ture, and Jimmy Lin*

University of Maryland Technical Report, HCIL-2010-23, October 2010.

**Personal Name Resolution in Email: A Heuristic Approach**

*Tamer Elsayed, Galileo Namata, Lise Getoor, and Douglas W. Oard*

University of Maryland Technical Report, TR-LAMP-150, March 2008.

**ATP: Autonomous Transport Protocol**

*Tamer Elsayed, Mohamed Hussein, Moustafa Youssef, Tamer Nadeem, Adel Youssef, and Liviu Iftode*

Tech. Rep. UMIACS-TR-2003-52 and CS-TR-4483, University of Maryland, May 2003.

**REFEREED ABSTRACTS**

**Effective Recommendation of Reviewers to Research Proposals**

*Nassma Mohandes, Qutaibah Malluhi, and Tamer Elsayed*

Qatar Foundation Annual Research Conference, Poster, 2014.

## RESEARCH GRANTS

---

### **Efficient and Scalable Evaluation for Searching Massive Arabic Social Media and Web Collections [as Lead-PI]**

Qatar National Research Fund (QNRF), National Priorities Research Program (NPRP) 7<sup>th</sup> Cycle, 3 years starting April 2015, \$883,847.

Joint with Matt Lease (University of Text, Austin, USA)

### **Scalable Analytics Engine for Big Graphs on the Cloud [as PI]**

Qatar National Research Fund (QNRF), National Priorities Research Program (NPRP) 7<sup>th</sup> Cycle, 3 years starting Dec. 2015, \$899,234.

Joint with Muhammad Hammoud (Carnegie Mellon University-Qatar, Qatar) and Ramy Melhem (University of Pittsburg, USA)

### **Answering Real-time Questions from Arabic Microblogs [as Lead-PI]**

Qatar National Research Fund (QNRF), National Priorities Research Program (NPRP) 6<sup>th</sup> Cycle, 3 years starting Feb 2014, \$1,045,414.

Joint with Doug Oard (University of Maryland, College Park, USA) and Mark Sanderson (RMIT, Australia)

### **Smart Shopping Assistant [as Sole PI]**

Qatar University internal student grant, 1 year starting Dec 2014, \$2,750.

### **Effective Retrieval and Filtering in Microblogs [as Sole PI]**

Qatar University internal student grant, 1 year starting June 2013, \$2,750.

### **Real-time Search in Twitter [as Sole PI]**

Qatar University internal start-up grant, 1 year starting April 2013, \$11,000.

### **Qcloud: Towards a Cloud Computing Infrastructure in Qatar to Target Regional Scientific Applications [as Co-PI]**

Qatar National Research Fund (QNRF), National Priorities Research Program (NPRP) 3<sup>rd</sup> Cycle (2012), 3 years starting Dec 2010, \$1,048,618.

Joint with Majd Sakr (Carnegie-Mellon in Qatar) and Qutaibah Malluhi (Qatar University)

## RESEARCH PROJECTS

---

### **Efficient and Scalable Evaluation for Searching Massive Arabic Social Media and Web Collections [as Lead-PI]**

*With Matt Lease, Assistant Professor, University of Texas - Austin, USA.*

### **Answering Real-time Questions from Arabic Microblogs [Feb 2014-Present]**

*With Douglas Oard, Professor, University of Maryland, College Park, USA.*

*With Mark Sanderson, Professor, RMIT, Australia.*

### **Query Performance Prediction in Microblog Search [Oct 2013 – Aug 2014]**

### **Qcloud: Cloud Computing at Qatar [Sep 2012- Sep 2013]**

*With Qutaibah Malluhi, Professor, Qatar University*

*With Majd Sakr, Associate Teaching Professor, Carnegie Mellon-Qatar*

### **Pick a Reviewer: A Recommender System for Grant Proposals [Sep 2012- May 2013]**

*With Qutaibah Malluhi, Professor, Qatar University*

### **Real-time Search in Twitter [June 2011 – Nov. 2011]**

*With Hany Ramadan, Assistant Professor, MCSE Division, KAUST*

- Designing and developing an effective and scalable real-time search system for twitter and participating in TREC 2011 micro-blog track for the first time at KAUST. Ranked 6<sup>th</sup> among 59 research teams.

### **iHadoop: MapReduce for Iterative Algorithms [Nov 2010 – Sep. 2011]**

*With Hany Ramadan, Assistant Professor, MCSE Division, KAUST*

- Proposing and developing a modified version of MapReduce that efficiently supports iterative algorithms, e.g., PageRank and K-means clustering.

### **Pseudo Test Collections [July 2010 – May 2011]**

*With Don Metzler, Computer Scientist, Information Sciences Institute, University of Southern California*

*With Jimmy Lin, Associate Professor, UMIACS, University of Maryland, College Park.*

- Leveraging the rich set of anchor text in the web to automatically generate a training set for learning-to-rank algorithms and a test collection for web search system evaluation.
- Cross-Lingual Pairwise Similarity in the Web [June 2010 – May 2011]**  
*With Jimmy Lin, Associate Professor, UMIACS, University of Maryland, College Park.*
- Designing and developing an efficient cross-lingual pairwise similarity algorithm that detects similar documents in two different languages.
- Ivory: A Scalable Open-source Search Engine [June 2009 – May 2011]**  
*With Jimmy Lin, Associate Professor, UMIACS, University of Maryland, College Park.*
- Developing a scalable search engine with end-to-end MapReduce implementation
  - Scaling near-duplicate detection algorithm to web-scale
  - Ranked 1<sup>st</sup> in Ad-hoc task of Web Track (Cat-B) in both TREC 2009 and TREC 2010
- Joint Resolution of Personal Names in Email Collections Using MapReduce [Feb 2008 – August 2009]**  
*With Douglas Oard, Professor, UMIACS, University of Maryland, College Park.*  
*With Jimmy Lin, Associate Professor, UMIACS, University of Maryland, College Park.*
- Designed a graph-based algorithm to jointly resolve all named-mentions in email collections
  - Designed and developed a distributed context expansion techniques for emails
  - Designed and developed an efficient pair-wise document similarity technique over large text collections using MapReduce
  - Built the largest and most-balanced test collection for the task of mention resolution in email
- JIKD: Joint Institute of Knowledge Discovery-Email Project [June 2005 – Feb. 2008]**  
*With Douglas Oard, Professor, UMIACS, University of Maryland College Park.*
- Designed and developed a search engine GUI for two archives of email: Enron and W3C.
  - Designed, developed, and evaluated a new model of identity for email archives.
  - Designed, developed, and evaluated an evidence-based heuristic approach for named mention resolution in email.
  - Designed, developed, and evaluated a probabilistic approach for named mention resolution in email using context expansion.
  - Designed, developed, and evaluated an algorithm for expert finding in Enterprise archives [as participation in TREC-2006 (Enterprise track, Expert Search task)]
- CASS: Context-Aware Search System [June 2004 – May 2005]**  
*With Douglas Oard, Professor, UMIACS, University of Maryland College Park.*
- Designed and implemented a novel formative evaluation strategy for adaptive topic tracking systems that accounts for the dynamic nature of those systems.
  - Designed, implemented, and evaluated an adaptive topic tracking algorithm based on n-gram language models [as a participation in TDT-2004 (Topic Tracking track)]
- Studying Facial Expressions as an Implicit Feedback in Information Retrieval Systems [Class Project, Oct. 2004 - Dec. 2004]**  
*With François Guimbretière, Assistant Professor, Dept of Computer Science, UMD.*
- Conducted (in a team of 2) a feasibility study on predicting document relevance from the user's facial expressions, which revealed that the detected facial expressions corresponding to relevant and non-relevant documents were distinguishable by a trained neural network classifier but many collected data are required to predict the relevance of future documents.
- Instance-Based Networking [May 2003 – May 2004]**  
*With Liviu Iftode, Associate Professor, Dept of Computer Science, Rutgers University.*
- Implemented and designed (in a team of 5) an extended version of a generic Content-Based Network (CBN) that acts as an overlay communication platform over which end-point entities communicate independently from their physical locations.
- Autonomous Transport Protocol [May 2003 – May 2004]**  
*With Liviu Iftode, Associate Professor, Dept of Computer Science, Rutgers University.*
- Design and implemented (in a team of 5) a transport protocol that allows mobile users to change networks and hosts seamlessly with continuous communication even if the user is not available for a short period of time.

**Distributed Object Lookup Service Over NICE [Class Project, Oct. 2002 - Dec. 2002]**

*With Bobby Bhattacharjee, Assistant Professor, Dept of Computer Science, UMD.*

- Designed and implemented (in a team of 3) an efficient object lookup service for NICE (a peer-to-peer *cooperative* framework for scalably implementing distributed applications over the Internet) group members by leveraging its good locality properties.

**Feedback Consolidation in Point-to-Multipoint Connections of ABR Service in ATM Networks.** [M.Sc. Thesis, Oct. 1998 - July 2001]

*With Nazih Elderini, Professor, Dept of Computer Science, Alexandria University, Egypt.*

- Designed, implemented, and evaluated an improved algorithm for feedback consolidation in point-to-multipoint ABR connections within ATM Networks. The algorithm combines benefits from the previous developed algorithms with reduced overhead. Results of the simulation experiments indicate that the proposed algorithm doesn't suffer from the consolidation noise, while exhibiting a fast transient response with accurate feedback.

**Human Face Recognition [B.Sc. Graduation Project, Oct. 1996 - July 1997]**

*With Mohamed Ismail, Professor, Dept of Computer Science, Alexandria University, Egypt.*

- Designed and implemented (in a team of 4), a near-real-time computer system that locates, tracks a subject's head, and then recognize the person by comparing characteristics of the face to those of known individuals. The system functioned by projecting face images onto a feature space that spans the significant variations among known face images.

**THESES SUPERVISION**

---

**Hind Al-Merekhi, M.Sc. [As advisor]**

Title: Building a Test Collection for Detecting Significant-Events in Arabic Tweets

Qatar University, QA

Graduated: Fall 2015

**Latifa AlMarri, M.Sc. [As advisor]**

Title: Ad-hoc Search in Microblogs using Hashtags

Qatar University, QA

Graduated: Spring 2015

**Maram Hasanain, M.Sc. [As advisor]**

Title: Query Performance Prediction in Microblog Search

Qatar University, QA

Graduated: Fall 2014

**Nassma Mohandes, M.Sc. [As a co-advisor]**

Title: Effective Recommendation of Reviewers to Research Proposals

Qatar University, QA

Graduated: Spring 2013

**Eslam Elnikety, M.Sc. [As a co-advisor]**

Title: iHadoop: Asynchronous Iterations for MapReduce

King Abdullah University of Science and Technology, KSA

Graduated: Summer 2011

**SOFTWARE RELEASES**

---

**Ivory, an open-source Hadoop toolkit for web-scale information retrieval research**

available at <http://ivory.cc>

- release 0.4 (2010/10/24)
- release 0.3 (2010/07/19)
- release 0.2 (2009/11/18)
- release 0.1 (2009/07/18)

## TEACHING EXPERIENCE

---

### **Qatar University**

#### *as an Instructor*

- CMPS 200: Computer Ethics (Spring 2016) [*Under-graduate*]
- CMPS 251: Object Oriented Programming (Spring 2016) [*Under-graduate*]
- CMPS 653: Big Data Analytics (Fall 2016) [*PhD*]
- CMPS 466: Information Retrieval (Fall 2016) [*Under-graduate*]
- CMPT 521: Information Retrieval (Spring 2015) [*Graduate*]
- CMPS 200: Computer Ethics (Spring 2015) [*Under-graduate*]
- CMPS 303: Data Structures (Fall 2014) [*Under-graduate*]
- CMPS 303: Data Structures (Spring 2014) [*Under-graduate*]
- GENG 107: Engineering Skills and Ethics (Spring 2014) [*Under-graduate*]
- CMPS 466: Information Retrieval (Fall 2013) [*Under-graduate*]
- CMPT 521: Information Retrieval (Spring 2013) [*Graduate*]
- GENG 106: Computer Programming (Spring 2012) [*Under-graduate*]
- GENG 106: Computer Programming (Fall 2012) [*Under-graduate*]

### **King Abdullah University of Science and Technology (KAUST), SA [Spring 2011 – Present]**

#### *as an Instructor*

- CS 238: Foundations of Information Retrieval (Summer 2011) [*Graduate*]

#### *as a Co-Instructor*

- CS 346: Large-Scale Computing (Spring 2011) [*Graduate*]

### **University of Maryland, College Park, US [Fall 2002-Spring 2004]**

#### *as a Teaching Assistant*

- CMSC 351: Algorithms (Spring 2004)
- CMSC 250: Discrete Structures (Fall 2003)
- CMSC 106: Introduction to C Programming (Fall 2002 and Spring 2003)

### **University of Alexandria, Egypt [Spring 1998-Spring 2002]**

#### *as a Teaching Assistant*

- Introduction to Computers
- Introduction to Programming
- Digital Design
- Data Structures
- Statistics
- Computer Networks and Communication
- Numerical Analysis
- Automatic Control

## OTHER WORK EXPERIENCE

---

### **Google, Mountain View, USA**

#### **Software Engineer Intern, Summer 2007**

**Starter project:** Internal-tool adoption: implemented a process that collects usage information of the tool's users by building a histogram of incoming/outgoing interactions.

**Main project:** Suggesting "Canned-Responses" to Customers

- Designed and implemented an infrastructure for pipelining, building ground truth, and preprocessing.
- Proposed different modeling approaches of providing automatic suggestions of "canned responses".

### **KME Computools, Alexandria, Egypt**

#### **Software Engineer, Fall 1997-Fall 1999**

Designed and implemented marketing and accounting applications using MFC, VC++ 5.0.

## GRADUATE COURSE WORK

---

Design and Analysis of Computer Algorithms, Advanced Computer Networks, Programming Language Technologies and Paradigms, Advanced Topics in Computer Systems, Advanced Algorithms, Neural Modeling, Natural Language Processing, Computer Graphics, Introduction to Artificial Intelligence, Advanced Introduction to Human-Computer Interaction, Cloud Computing: Web-Scale Information Processing Applications.

## TALKS

---

- What Questions Do Journalists Ask on Twitter? The Social Media in the Newsroom Workshop (SMNews' 16) at International Conference on Web and Social Media (ICWSM). May 2016.
- bigIR: When Big Data meets Information Retrieval. Big Data Management workshop organized at QU. Apr 2016.
- MapReduce 101. Guest Lecture in "CMPT 523 Distributed Systems" graduate class, QU, Mar 2016.
- Detecting Automatically-Generated Arabic Tweets, AIRS 2015, Dec 2015.
- Researcher 101: Issues You Would Like to Know, QU CSE Seminar, Feb 2015.
- QU @ TREC-2014: Online Clustering for Tweet Timeline Generation, TREC Conference, Gaithersburg, MD, USA, Nov 2014.
- Pairwise Document Similarity in Large Collections, Guest Lecture in "CMPT 571 Advanced Algorithms" graduate class, QU CSE Seminar, Oct 2014.
- Research Issues You Must Know, Research 101 Workshop, QU Seminar, Oct 2014.
- How to Read a Research Paper, Research 101 Workshop, QU Seminar, Oct 2014.
- Research? A Gentle Introduction, Research 101 Workshop, QU Seminar, Oct 2014.
- MapReduce 101. Guest Lecture in "CMPT 507 Advanced Operating Systems" graduate class, QU, May 2014.
- Researcher 101: Issues You Would Like to Know, QU CSE Seminar, April 2014.
- IR in a Nutshell: Applications, Research, and Challenges, QU CSE Seminar, February 2014.
- TREC-2013: QU Taking the "Fast Lane" to IR Research, QU CSE Seminar, Oct 2013.
- LaTeX: Writing Professional Scientific Documents, OFID EduTech Day 2013, QU, June 2013.
- How to Read a Research Paper, QU CSE Seminar, May 2013.
- IR in a Nutshell: Applications, Research, and Challenges, QU CSE Seminar, February 2013.
- On Large-Scale Retrieval Tasks with Ivory and MapReduce, QU CSE Seminar, November 2012.
- On Large-Scale Retrieval with Ivory, Microsoft Research at Cambridge, October 2011.
- Scalable Identity Resolution in Email Collections with MapReduce. Division Seminar, CMSE Division, KAUST, November 2010.
- Scalable Identity Resolution in Email Collections Using MapReduce. Guest Lecture in "Data-Intensive Information Processing Applications" Class, University of Maryland, College Park, MD, April 2010.
- Identity Resolution in Email Collections. CLIP Lab Colloquium, University of Maryland, College Park, MD, February 2009.
- No, Not That PMI: Creating Search Technology for E-Discovery. Joint talk with Jason Baron, Douglas W. Oard, and Lidian Wang. iSchool Colloquium, University of Maryland, College Park, MD, October 2008.
- Computing Pairwise Document Similarity in Large Collections: A MapReduce Perspective, **Guest Lecture** in "Introduction to Cloud Computing" Class, University of Maryland, College Park, MD, October 2008.
- Resolving Personal Names in Email using Context Expansion, ACL Conference, Columbus, OH, June 2008.
- Pairwise Document Similarity in Large Collection with MapReduce, ACL Conference, Columbus, OH, June 2008.

- Using MapReduce for Scalable Coreference Resolution, HLT COE at JHU, Baltimore, MD, June 2008.
- Ivory, “Scaling Up Language Technologies” Mini-workshop, CMU, Pittsburgh, PA, May 2008.
- The Enron and W3C Collections, ICAIL DESI Workshop, Stanford, CA, June 2007.
- Modeling Identity in Archival Collections of Email: A Preliminary Study, CEAS Conference, Mountain View, CA, July 2006.

## **COMPUTER SKILLS**

---

*Programming Languages:* C, C++, VC++, Assembly, Pascal, BASIC, and Java.

*Operating Systems:* DOS, Microsoft Windows, Unix

*Open Source:* Hadoop, Storm, Lucene, JavaMail, Lingpipe, JDOM, and WEKA

*Database Systems:* Ms FoxPro, and MS Access.

## **SCHOLARLY ACTIVITIES**

---

Reviewer, IEEE TKDE, 2016

Reviewer, IP&M 2016

Reviewer, JASIST 2016

TPC member, SIGIR-Short Papers 2016

Reviewer, SemEval 2016

TPC member, The 2nd Workshop on Arabic Corpora and Processing Tools, LREC 2016.

TPC member, NAACL 2015

Reviewer, IEEE TKDE, 2014

TPC member and IR track co-chair, AICCSA 2014

Reviewer, Foundations and Trends in Information Retrieval, 2013

TPC member, ICCIT 2013

Reviewer, IEEE TKDE, 2012

Reviewer, ACM SIGIR, 2011

Reviewer, Journal of Information Retrieval (IRJ), 2010

Reviewer, ACM Transactions on Information Systems (TOIS), 2009

Reviewer, ACM SIGIR, 2008

Reviewer, ACM SIGIR, 2006

Member, ACM SIGIR

## **OTHER ACTIVITIES**

---

Google Ambassador at University of Maryland-College Park, 2007-2008

Member, Syndicate of Egyptian Engineers