tamer.elsayedaly@kaust.edu.sa 4700 KAUST, Thuwal 23955-6900, Jeddah, Saudi Arabia
http://www.cs.umd.edu/~telsayed +966 2 808 0322

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

Post-Doctoral Fellow, September 2010 – Present
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, CLIP Lab
University of Maryland-College Park, MD, US

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

RESEARCH INTERESTS

GENERAL INTERESTS: INFORMATION RETRIEVAL & CLOUD COMPUTING

- Cloud Computing with emphasis on large-scale text analysis
- Designing and evaluation of search engines
- Searching informal media (e.g., emails)

PHD TOPIC: IDENTITY RESOLUTION IN EMAIL COLLECTIONS

Resolving identity of people mentioned in Enterprise email archives where the searcher is not a participant.

Some of the inherent research problems are:

- Building computational model of identity of participants in email collections
- Context expansion of email messages
- Resolution of single mentions in the email body
- Scalable Joint resolution of *all* mentions in the whole collection using MapReduce
- Development of the largest test collection available for research for this task

PUBLICATIONS

JOURNAL PUBLICATIONS

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

Tamer Elsayed, Ferhan Ture, and Jimmy Lin

Submitted to Information Retrieval Journal

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*), accepted in October 2009.

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

No Free Lunch: When Locality-Sensitive Hashing Performs Worse than Brute Force for Cross-lingual Pairwise Similarity

Ferhan Ture, Tamer Elsayed, and Jimmy Lin

Full paper submitted to SIAM Conference on Data Mining (SDM) 2011.

Arabic Cross-Document Coreference Detection

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

The 47th 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 46th 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 46th 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 3rd 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

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.

UNREFEREED PAPERS

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, November 2010.

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.

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, November 2009.

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.

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, 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, November 2004.

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.

RESEARCH EXPERIENCE

iHadoop: MapReduce for Iterative Algorithms [November 2010 – Present]

With Hany Ramadan, Assistant Professor, MCSE Division, KAUST

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

Ivory: A Scalable Open-source Search Engine [June 2009 – Present]

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

www.ivory.cc

- Developing a scalable search engine with end-to-end MapReduce implementation
- Scaling near-duplicate detection algorithm to web-scale
- Designing and developing a cross-lingual pairwise similarity algorithm
- Ranked 1st 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.

TEACHING EXPERIENCE (AS A TEACHING ASSISTANT)

University of Maryland, College Park: Fall 2002-Spring 2004

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

University of Alexandria, Egypt: Spring 1998-Spring 2002

- 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

- 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 Lidan 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, Lucene, JavaMail, Lingpipe, JDOM, and WEKA
Database Systems: Ms FoxPro, and MS Access.

ACTIVITIES

Google Ambassador at University of Maryland-College Park, 2007-2008
Student Member, ACM SIGIR
Reviewer, Journal of Information Retrieval (IRJ), 2010
Reviewer, ACM Transactions on Information Systems (TOIS), 2009
Reviewer, ACM SIGIR, 2008
Reviewer, ACM SIGIR, 2006
Member, Syndicate of Egyptian Engineers