

Murat Ali BAYIR
Research Assistant
Ubiquitous Computing Laboratory
Computer Science and Engineering
SUNY at Buffalo

Personal Information

Address 210 Furnas Hall, Amherst, Buffalo, NY 14260
E-mail mbayir@cse.buffalo.edu
Mobile Telephone +1 (716) 4184060
www www.cse.buffalo.edu/~mbayir

Education

- 2007 — now University at Buffalo, NY, USA
- Ph.D. in Department of Computer Science and Engineering
 - **Supervisor:** Prof. Murat Demirbas
- 2006 — 2007 Middle East Technical University, Ankara, Turkey.
- PhD student for one year.
- 2003 — 2006 Middle East Technical University, Ankara, Turkey.
- MS in Department of Computer Engineering
 - **Thesis Title:** A new Reactive Method For Processing Web Usage Data
 - **Supervisor:** Prof. I. Hakki Toroslu
- 2001 — 2004 Middle East Technical University, Ankara, Turkey.
- Minor degree in Department of Mathematics
- 1999 — 2003 Middle East Technical University, Ankara, Turkey.
- BS in Department of Computer Engineering
 - **Thesis Project:** Web Based Banking Automation System
 - **Supervisor:** Prof. Ahmet Cosar
- 1998 — 1999 Middle East Technical University, Ankara, Turkey.
- One Year Language Preperation in Department of Basic English
- 1994 — 1998 Mersin 19 Mayis Anatolian High School (English), Mersin, Turkey.
- High School Degree, Natural Science Division.
-

Research Areas:

- Mining Human Mobility and its application to Wireless Sensor Networks
- Data Mining, Web Mining
- Machine Learning

Courses Taken at UB during PhD Study

- Algorithms
 - Machine Learning
 - Wireless Sensor Networks
 - Seminar in Wireless Sensor Networks
-

Publications

- **Under Submissions (* means publications related to PhD study)**

1. Murat Ali Bayir, Murat Demirbas, Nathan Eagle, "Mobility Profiler: A Framework for Discovering Mobile User Profiles". (*)
2. Ismail H. Toroslu, Murat Ali Bayir, " Inserting Link Information into Web User Sessions for Frequent Navigation Pattern Discovery".
3. Ismail H. Toroslu, Murat Ali Bayir, "Graph Theoretic Approach for Session Reconstruction Problem".
4. Murat Demirbas, Atri Rudra, Carole Rudra, Murat Ali Bayir, "iMAP: Indirect Measurement of Air Pollution with Cellphones". (*)

- **Journals**

5. Murat Ali Bayir, Dogacan Guney, Tolga Can "Integration of Topological Measures for Eliminating Non-Specific Interactions in Protein Interaction Networks". (In press) Discrete Applied Mathematics Journal, Elsevier, Special issue on Biological Networks. 2007.
6. Murat Ali Bayir, Ismail H. Toroslu, Ahmet Cosar, "Genetic Algorithm for Multiple Query Optimization problem". IEEE Transactions on Systems, Mans and Cybernetics. Part C. Volume:37, Issue:1 page(s): 147-153.

- **Conference and Workshop Papers**

7. Murat Ali Bayir, Ismail H. Toroslu, Ahmet Cosar, Guven Fidan " Discovering More Accurate Frequent Web Usage Patterns". Text-Mining & Link-Analysis 2007 at IJCAI.
8. Murat Ali Bayir, Ismail H. Toroslu, Ahmet Cosar, "A New Approach for Reactive Web Usage Data Processing," *icdew*, p. 44, 22nd International Conference on Data Engineering Workshops (ICDEW'06), 2006.
9. Murat Ali Bayir, "BASKIN: Small Scale Combat Simulation System", ITEC 2006, European Conference on Simulation and Training Systems.
10. Murat Ali Bayir, Ismail H. Toroslu, Ahmet Cosar, "A Performance Comparison of Pattern Discovery Methods on Web Log Data", AICCSA-06, The 4th ACS/IEEE International Conference on Computer Systems and Applications, 2006.

- **MS Thesis and Other Publications**

11. Murat Ali Bayir, "A New Reactive Method For Processing Web Usage Data", MS Thesis, Middle East Technical University, June 2006.
12. Murat Ali Bayir, "Introduction to Web Programming", online course book for web Programming course of IDEA of CENG-METU, 2003. (in Turkish)
13. Murat Ali Bayir, Ismail H. Toroslu, Ahmet Cosar, Guven Fidan " Discovering More Accurate Frequent Web Usage Patterns". CoRR abs/0804.1409, Cornell Digital Library.

Honors And Awards

03/2006 – 05/2007	Accepted to PhD program in Computer Science in the following Universities. For Fall 2007 <ul style="list-style-type: none">• SUNY at Buffalo, USA, (With Reseach Asistant Support)• Case Western Reserve University, USA, (With Reseach Asistant Support)• SUNY at Binghamton USA• University of Glasgow, UK, (With Reseach Asistant Support) For Winter 2007 <ul style="list-style-type: none">• ETH Zurich, Switzerland, (With Reseach Asistant Support) For Fall 2006 <ul style="list-style-type: none">• Northeastern University, USA, (With Reseach Asistant Support)• Arizona State University, USA, (With Reseach Asistant Support)• Boston University, USA• Middle East Technical University, Turkey
11/2006	Promoted to Project Coordinator of Bilgi.com Search Enginee Project.
07/2004	Winner of WeberDev PHP code contest with source code on server side image processing with imagemagick in PHP.
11/2003	Invited Speaker on Session Management in PHP, Speech is given to the Computer Center Members of Turkish Environment Ministry.
06/2003	Graduated as a honor student from Computer Engineering Department, METU.
07/1998	Ranked 582 th in university entering exam among 1.4 million candidates.

Memberships and organizations participated in

- IEEE Student Member
 - Organization Committee Member, ICDE 07. IEEE 23 rd International Conference on Data Engineering (web: www.icde2007.org)
 - Participated in AB-07, National Academic Conference on Information Systems.
 - Participated in IST-2006 Event, European Meeting in Information and Telecommunication Research Society. Give talks on Web Usage Mining and AGMLAB's Research Vision.
 - Baskin Product Presentation, USMOS 05, National Modeling and Simulation Symposium of Turkey.
-

Research Projects

Date - Place	Project
2006 - AGMLAB Inc Senior Software Engineer	<p style="text-align: center;">Topic Based Focused Crawler</p> <p>Implemented focused crawler for fetching topic-specific web pages in www. An initial word set is modeled as unit vector for representing particular domain. Each web site is represented by a topic vector containing frequency of each element in the unit vector of corresponding domain. Under the vector space modeling, the cosine of the angle between unit vector of that domain and topic vector of corresponding page is calculated. This value stands for similarity of web page to corresponding domain. This system uses a similarity threshold q ($0 < q < 1$) for storing current web page as a related one for given domain. If similarity is greater than q, corresponding web page is marked as related to that domain; other web pages not satisfying the rule are pruned. Web crawling process continues depth by depth with pruning search space in this manner until enough data is collected.</p>
2006 - AGMLAB Inc Senior Software Engineer	<p style="text-align: center;">Design and Implementation of Page Importance Calculation Methods on Web Graph</p> <p>Implemented various page importance calculation methods including OPIC, ordinary Page Rank, edge weighted page Rank and two layered Page Rank. An Evaluation framework is designed for testing performance of these page importance methods. In this context, Tao Sequence Similarity Test and DMOZ Coverage Tests are Implemented. Comparison of page importance calculation methods is performed with respect to top internet sites in more than 15 categories in Turkish Web.</p>
2004/2005 - METU As a Graduate Student	<p style="text-align: center;">A New Session Reconstruction Method for Web Usage Mining</p> <p>In this project, a new session reconstruction method called Smart-SRA is developed for session construction phase of web usage mining. Unlike the simple time based approaches, the session reconstruction problem is modeled as a graph problem the main idea of which is to insert the link information into web user session during session construction. The Approach is novel and we have explored sessions satisfying increasing time stamp order and link topology rules from web usage data. We have also shown that the sessions generated by Smart-SRA are more accurate than the sessions generated by previous heuristics.</p>
2004/2005 - METU As a Graduate Student	<p style="text-align: center;">A Web Agent Simulator, Simulating Navigation Request of Web User</p> <p>Developed an agent simulator that generates web agent requests by simulating an actual web user. This system is an improved model of google's page-ranks random surfer. In addition to behaviors in random surfer model, We have added new user behaviours like backward movements and session termination. This system is used for evaluation of alternative session reconstruction heuristics since processing server side access logs cannot provide entire actual user requests. Especially the sessions containing access requests served from a client's and/or proxy's local cache cannot be accurately determined. However, we can compare the alternative session construction heuristics since this web agent simulator gives actual user behavior of simulated web users on client-side.</p>
2004 - METU	<p style="text-align: center;">Author Recognition From Text Context</p> <p>The aim of the project is that writer of any text is going to be recognized with the help of previous training data. The decision tree approach is going to be used as pattern recognition</p>

As a Graduate Student Course Project	technique. Training data contains previous texts of corresponding writers. Any author is represented by feature vector. We have used commonly used patterns of author and LZW compression algorithm for calculating distance of current text with each author. After that, each distance value is given as an attribute in our decision tree approach.
2003 - METU As a Graduate Student	Evolutionary Computing Methods for MQO Problem
	MQO (Multiple Query Optimization) problem is a well-known database research problem, and the database community has studied it since 1980s. The goal of MQO is to reduce the execution cost of a set of queries by performing their common tasks only once. In this project, MQO problem is modeled for Genetic Algorithms. In the context of MQO, a chromosome corresponds to a solution instance for the set of queries of the MQO problem. In a chromosome, each gene of a chromosome represents a plan to the corresponding query. Under this modeling, it is shown that MQO is also very suitable for genetic operations. Crossover and mutation operations can easily be defined to produce new valid solution instances. It is shown that By using GA, it is possible to obtain total execution times for MQO very close to optimum execution value in polynomial time.

Professional Experience, Selected Best Projects

Project Number	Date	Developed Projects and Activities (Where position was taken)	Project Area
1	2006 - 2007	<p>MANAGEMENT ACTIVITIES (AGMLAB Inc):</p> <ul style="list-style-type: none"> • Project coordinator activities of bilgi.com for 8 months. • Participated in both project management and Turkish NSF grant proposal tasks. • Prepared development reports of Bilgi.com project for submitting TUBITAK (NSF of Turkey). • Participated in project proposal preparation of three European FP-7 Program Research Projects, FRICOC, INQUSON and MIRACLE. <p>BILGI, Turkish Search Engine Project (AGMLAB Inc). (www.bilgi.com)</p> <p>Project Description: Bilgi.com is the first Turkish search engine to consider the Turkish morphological structure and to present Turkish web content with high technology. Bilgi.com reaches whole Turkish content, index it periodically and keep it up to date. The project has been supported by Scientific & Technological Research Council of Turkey (TUBITAK) and Technology Development Foundation of Turkey (TTGV).</p> <p>Activities During the Project</p> <ul style="list-style-type: none"> • Participated in both development and improvement of crawler module. Designed both site Specific and focused crawler. Implemented OPIC scoring system for sorting urls during crawl operation. This system is used for selecting which urls will be fetched. • Participated in both development and improvement of search quality system. Implemented trustable version of Page Rank Algorithm for ranking during indexing phase. • Implemented link based semi-automated SPAM elimination system over whole Turkish Web Graph. The software found Potential Spam hosts in Turkish web graph. This software also used for reducing size of Search space in web graph for manual spam page detection. 	Information Systems

2	2005-2006	<p>BASKIN, the project being developed by cooperation's of Havelsan, Modsim and Turkish Armed Forces. (HAVELSAN Inc.)</p> <p>Project Description: BASKIN is the first national simulation system based on high-level architecture by using agent based programming. It is developed for simulating commando and small defense operations by using recent modeling and simulation technologies.</p> <p>Activities During the Project</p> <ul style="list-style-type: none"> • Resolved approximately 60 fatal bugs of the system and deployed Baskin Beta System in Computer Laboratuvary of Turkish Armed Forces. Gives training to 30 soldiers. • Participated in software maintenance of BASKIN for one year. Solved fatal bugs of BASKIN Simulation System. Published V1.0 Beta Version of BASKIN. • Participated in Reverse engineering activities and prepared software design, requirements and specification of system proper to CMM 3 standards. 	Simulation And Training Systems
3	2004-2005	<p>(UYAP) National Justice Sytems which is a part of National E-revolution project of Turkey (HAVELSAN Inc.)</p> <p>Project Description: UYAP is the one of the biggest information system project in Turkey. It is the part of e-revolution project of Turkey such that it interacts with other systems of e-government. It is developed for decreasing time for paper works in national justice system.</p> <p>Activities During the Project</p> <ul style="list-style-type: none"> • Participated in analysis, design of Yargitay Module. • Implemented investigation, searching and voting subsystems of Yargitay Module. 	Information Systems
4	2003-2004	<p>Turk Telecom Transnet Automation System (Ideal Technology)</p> <p>Project Description: Transnet Automation system keeps the local information of network nodes in Turk-Telecom System in Turkey. It is the part of e-revolution project of Turkey such that it interacts with other systems of e-government.</p> <p>Activities During the Project</p> <ul style="list-style-type: none"> • Participated in requirement analysis, software design and implementation of the Yargitay Module. • Implemented investigation, searching and voting subsystems of Yargitay Module. 	Information Systems
5	2003	<p>Hipfoto Communication Portal for Yalçınlar A.Ş. (Ideal Technology) http://hipfoto.com</p> <p>Project Description: Hipfoto is the drupal based high technology portal based system. It is developed for photograph sharing among portal communitiy, image processing and editing online. It has an advanced feautes like server side image processing when compared to similar projects.</p> <p>Activities During the Project</p> <ul style="list-style-type: none"> • Participated in requirement analysis, software design and 	Information Systems

		implementation of the project. <ul style="list-style-type: none"> Implemented file storage, photo editing, photo sharing and admin sub modules of the system. 	
--	--	--	--

Professional Experience, Company Based

Company (Sector)	Date	Position
AGM Lab Inc. (Information Systems)	(01.05.2006 – 01.08.2007)	Project Coordinator & Research Engineer
Havelsan Simulator Systems (Defence Industry)	(01.05.2005 – 01.05.2006)	Senior Software Engineer
Havelsan Sivil Bilisim Dir. (Information Systems)	(01.04.2004 - 01.05.2005)	Software Engineer
Ideal Technology (Information Systems)	(15.06.2003 - 01.04.2004)	Software Engineer
METU-IDEA (Remote Education)	(01.11.2002 – 01.06.2003)	Student Assistant

Computer Skills

Operating Systems	Experience/Knowledge	Project No referred Above
Linux	Advanced	Not Applicable
Windows (2k-XP-2003server)	Advanced	Not Applicable
Web Servers	Experience/Knowledge	Project No referred Above
Apache Web Server	Administrator Level	1
Apache Tomcat	Administrator Level	1, 4
Portal Systems	Experience/Knowledge	Project No referred Above
Microsoft SPS Portal Services	Administrator Level	2
Drupal	Administrator Level	5
Development Tools	Experience/Knowledge	Project No referred Above
Eclipse	Advanced	1
IBM WSAD	Advanced	3
Microsoft Visual C++	Advanced	2
Intel. IDEA	Advanced	4
Zend Studio	Advanced	5
Configuration Mang. Tools	Experience/Knowledge	Project No referred Above
CVSNt-Server	Administrator Level	2
WinCVS-Client	Administrator Level	2
CVS linux	Administrator Level	4
Svn	Administrator Level	1
Technologies	Experience/Knowledge	Project No referred Above
Java	Advanced	1, 3, 4
C	Advanced	2
PHP	Advanced	5
C++	Mid-Level	2
JSP/Servlet	Beginner	4
Lisp	Beginner	2
Database Man. Systems	Experience/Knowledge	Project No referred Above
MySQL	Administrator Level	1, 2, 4, 5

Oracle API	Mid-Level Experience/Knowledge	3 Project No referred Above
Apache Nutch Search Engine API	Advanced	1
Apache Hadoop Map Reduce Distributed Processing API	Advanced	1
Apache Hadoop Distributed File System API	Advanced	1
SourceForge HTML Parser, Neco HTML Parser	Advanced	1
Apache Lucene Indexer API	Mid-Level	1
Win32, MFC	Mid-Level	2
Open GL	Mid-Level	2

Given Courses and Seminars

- **February 2007**, www.bilgi.com search engine exhibition and presentation, AB 2007, National Academic Conference on Information Systems.
- **November 2006**, presented Web Usage Mining Applications and Concepts, IST 2006, European Conference for FP7 Networking and Research.
- **November 2006**, presented AGMLAB's Research and FP7 Vision, IST 2006, European Conference for FP7 Networking and Research.
- **July-Agust 2005** BASKIN Administrator Level User Education to Special Force of Turkey.
- **June 2005**, BASKIN product exhibition and presentation, USMOS 2005 National Conference of Turkey on Simulation and modeling Systems.
- **November 2003** Session Management in PHP, to the Computer Center members of Turkish Environment Ministry.
- **2002-2003** Introduction to Web Programming, Course of METU-IDEA

Activities

- Reviewer for IEEE Transactions on Evolutionary Computing Journal
- External Reviewer for International Conference on Principles of Distributed Systems, OPODIS 07.
- External Reviewer for International Conference on Communication ICC 07.
- External Reviewer for WOWMOM 08.