TAMER KAHVECI

Department of Computer Science,
University of California,
Santa Barbara, CA, 93106

Tel : (805) 893-4276
Fax : (805) 893-8553
Email :tamer@cs.ucsb.edu
http://www.cs.ucsb.edu/~tamer

RESEARCH INTERESTS

Databases: Multimedia databases, biological databases, text databases, spatial databases, distributed databases, replication, caching.

Bioinformatics: Pattern matching, genome-wide analysis, comparative genomics, structural genomics, shotgun sequence assembly, motif finding.

EDUCATION

1999-Present: **University of California, Santa Barbara,**
Ph.D. Candidate, Computer Science.

1997-1999: **Bilkent University, Ankara, Turkey,**
M.S. Student, Computer Science.

1993-1997: **Bilkent University, Ankara, Turkey,**
B.S., Computer Science

RESEARCH & WORK EXPERIENCE

2000-Present: **Research Assistant, University of California, Santa Barbara,**
Worked on multimedia databases, sequence databases, and bioinformatics. Developed a number of techniques to answer similarity queries for sequence databases. These techniques lead towards efficient search tools for biological and time series. Also developed new methods which allow biological search tools to use memory efficiently and to run interactively. Participated in building a web server that enables complex biological queries.
Supervisor : Ambuj Singh

Summer 1999: **Software Engineer, at Bilkent University Computer Center, Ankara, Turkey,**
Designed and developed a new database management system for Bilkent University, Private Bilkent High School, and the owner company personnel system overcoming the Y2K problem.

August 1995: **Researcher at The Scientific and Technical Research Council of Turkey, Ankara**
Worked on performance improvements using ATM (Asynchronous Transfer Mode) for networks.

TEACHING EXPERIENCE

1999-2000: Teaching Assistant, University of California, Santa Barbara,
Undergrad Courses: Foundations of Scientific Computing, Machine Intelligence, Introduction to Formal Languages.
Grad Courses: Indexing in Multimedia Databases.
Graded homeworks, midterms, and finals. Helped prepare course material.
Managed office hours and discussions regularly, maintained course web page.

1997-1999: Teaching Assistant, at Bilkent University, Ankara, Turkey
Courses: Object Oriented Programming, Numerical Methods, Introduction to Programming and Algorithms, Combinatorics and Graph Theory.
Graded homeworks, midterms, and finals. Managed office hours and regularly.
Taught classes on instructors absence.

1993-1997: Trainer and Mentor of Turkish Mathematical Olympiad Team, Turkey
Helped students in Turkish Mathematical Olympiad team prepare for International Mathematical Olympiads for High School students.

REFEREED PUBLICATIONS

(available at http://www.cs.ucsb.edu/~tamer/publications.html)

- Tamer Kahveci, Vebjorn Ljosa, Ambuj K. Singh,
Speeding up Whole-Genome Alignment by Indexing Frequency Vectors,
to appear in **Bioinformatics**
- Orhan Camoglu, Tamer Kahveci, Ambuj K. Singh,
Towards Index-based Similarity Search for Protein Structure Databases,
to appear in **Journal of Bioinformatics and Computational Biology (JBCB)**, 2003 (invited paper).
- Tamer Kahveci, Ambuj K. Singh,
Optimizing Similarity Search for Arbitrary Length Time Series Queries,
to appear in **IEEE Transactions on Knowledge and Data Engineering (TKDE)**
- Arnab Bhattacharya, Tamer Kahveci, Ambuj K. Singh
Motif Finding and Multiple Alignment through Vector-Space Embedding of Protein Sequences,
to appear in **International Conference on Research in Computational Molecular Biology (RECOMB)**, 2004 (poster paper), San Diego, CA.
- Arnab Bhattacharya, Tolga Can, Tamer Kahveci, Ambuj K. Singh, Yuan-Fang Wang
"ProGreSS: Simultaneous Searching of Protein Databases by Sequence and Structure",
Pacific Symposium on Biocomputing (PSB), 2004, Big Island, Hawaii.
- Orhan Camoglu, Tamer Kahveci, Ambuj K. Singh,
Towards Index-based Similarity Search for Protein Structure Databases,
Computational Systems Bioinformatics (CSB), 2003, Stanford, CA.
- Tamer Kahveci, Ambuj K. Singh,
Genome on Demand: Interactive Substring Searching,
Computational Systems Bioinformatics (CSB), 2003 (poster paper), Stanford, CA.
- Orhan Camoglu, Tamer Kahveci, Ambuj K. Singh,
PSI: Indexing Protein Structures for Fast Similarity Search,
International Conference on Intelligent Systems for Molecular Biology (ISMB), 2003 (also in Bioinformatics Journal), Brisbane, Australia.

- Tamer Kahveci, Christian Lang, Ambuj K. Singh,
  Joining Massive High-Dimensional Databases,
  International Conference on Data Engineering (ICDE), 2003, Bangalore, India
- Tamer Kahveci, Ambuj K. Singh,
  Fast Alignment of Large Genome Databases,
  International Conference on Data Engineering (ICDE), 2003 (demo paper), Bangalore, India
- Tamer Kahveci, Ambuj K. Singh,
  MAP: Searching Large Genome Databases,
  Pacific Symposium on Bioinformatics (PSB), 2003, January, Kaua'i, Hawaii
- Tamer Kahveci, Ambuj K. Singh, Aliekber Gurel,
  Similarity Searching for Multi-Attribute Sequences,
  International Conference on Scientific and Statistical Database Management (SSDBM), 2002, Edinburgh, Scotland.
- Tamer Kahveci, Ambuj K. Singh, Aliekber Gurel,
  An Efficient Index Structure for Shift and Scale Invariant Search for Multi-attribute Time Series Data,
  International Conference on Data Engineering (ICDE), 2002 (poster paper), San Jose, CA.
- Tamer Kahveci, Ambuj K. Singh,
  An Efficient Index Structure for String Databases,
  International Conference on Very Large Databases (VLDB), 2001, Roma, Italy.
- Tamer Kahveci, Ambuj K. Singh,
  Variable Length Queries for Time Series Data,
  International Conference on Data Engineering (ICDE), 2001, Heidelberg, Germany.
- Tuba Yavuz-Kahveci, Tamer Kahveci, Ambuj Singh,
  Buffering of Index Structures,
  The International Society for Optical Engineering (SPIE) - Internet Multimedia Management Systems, 2000, Boston.

BOOK CHAPTERS

- Tamer Kahveci, Ambuj K. Singh,
  Index Structures for Approximate Matching in Sequence Databases
  to appear in "Handbook on Computational Biology", editor Srinivas Aluru, CRC Press

---

PATENTS

- Tamer Kahveci and Ambuj K. Singh, "Frequency Vectors for Accelerating String Searches" (US patent pending).

SELECTED TALKS

March 2003: ICDE, Bangalore, India, "Joining Massive High-Dimensional Databases".

January 2003: Invited talk in PSB, Kauai, "Challenges in Large Scale Genome Alignment".

October 2002: Seminar on Bioinformatics at UCSB, "Vector Alignment Search Tool".

July 2002: SSDBM, Edinburgh, Scotland, "Similarity Searching for Multi-Attribute Sequences".

February 2002: Seminar on Data Management Issues in Digital Libraries at UCSB, "K2/Kleisli and GUS: Experiments in integrated access to genomic data sources, by S. B. Davidson, J. Crabtree, B. P. Brunk, J. Schug, V. Tannen, G. C. Overton, and C. J. Stoeckert, Jr. appeared in IBM Systems Journal, Volume 40, Number 2, 2001".

February 2002: Seminar on Bioinformatics at UCSB, "FLASH: Fast Look-up Algorithm for String Homology, by A. Califano and I. Rigoutsos, appeared in CVPR, 1993".

Sept. 2001: VLDB, Roma, Italy, "An Efficient Index Structure for String Databases".

April 2001: ICDE, Heidelberg, Germany, "Variable Length Queries for Time Series Data".

March 2001: Thesis proposal at UCSB, "Sequence Databases".

October 2000: SPIE, Boston, "Buffering of Index Structures".

SERVICES

- **Refereed** for Bioinformatics, VLDB 2001, VLDB 2002, CIKM 2002, TODS, Workshop on Bioinformatics 2003 held in conjunction with ICDE 2003, ICDE 2004, EDBT 2004.
- **Student volunteer** at DIGIVATIONS 2000, SIGMOD 2001.
- **Graduate Admissions Committee** at Computer Science Department of UCSB (2002-2003).
- **Webmaster** of the Workshop on Bioinformatics held in conjunction with ICDE 2003.
- **Student member** of IEEE.

HONORS AND AWARDS

2000-2003: Presidents Work Study research grant.

2003: National Library of Medicine award for the Pacific Symposium on Biocomputing (PSB) 2003, January 3-7, Lihue, Hawaii.

2001: Graduate Division Travel Grant awarded by University of California, Santa Barbara.

1999: Scholarship for the Ph.D. program of the Computer Science Department of UCSB.

1997: Scholarship for the M.S. program of the Computer Science Department of Bilkent University.

1993: Scholarship for the B.S. program of the Computer Science Department of Bilkent University.

1993: Silver medal in 34th International Mathematical Olympiads among 413 competitors from 74 countries.

1993: Member of the Turkish National Mathematical Olympiads team in the 34th International Mathematical Olympiads.

1990: 1st rank in the chess tournament among high schools organized by the Ministry of National Education of Turkey.

SELECTED PROJECTS

- **Bioinformatics Web-Server:**
  Participated in building a bioinformatics web server. This server can answer a wide variety of biological queries. Some of these queries are:

  1. Searching a pattern similar to a given query pattern interactively.
  2. Aligning huge sequences, such as complete genome.
  3. Protein structure similarity search.
  4. Similarity search based on expression arrays and genomic pathways.

  The web server comprises well known search tools such as BLAST as well as tools that we developed. The server contains biological data obtained from various well known databases. The user can either perform queries on these databases or construct a personal database by uploading new data.
- **Improving Throughput of Distributed Systems:**
  Considered the problem of improving the throughput of distributed systems when some of the clients can be disconnected for some period of time. Clients are allowed to perform transactions while they are disconnected. In disconnected phase, some of the transactions are committed and some of the transactions are put in a pending list. When a disconnected user is connected again, the sequence of pending transactions performed by that client are merged to the other sequences of pending transactions. Proposed and implemented a sharing-based reordering scheme to increase the amount of committed transactions in the disconnected phase, and decrease the amount of rollbacks in the merge phase.
- **Searching Motion Similarity in MPEG Files:**
  Designed and implemented a Java program that extracts motion vectors of the macro blocks from MPEG files, performs DFT and applies Shift Eliminated Transformation in frequency domain to find the minimal Euclidean distance under the set of all possible shiftings and scalings using the R* tree as the index structure.
- **Personnel Database Project:**
  Designed and implemented a Delphi program that allows all the facilities of a business information system such as inserting, deleting, modifying, searches, monthly or yearly payment computation, interactive or automatic list production, template definition, information sharing through the network, adaptability to changing laws and many other features. This project is developed to meet the needs of Bilkent University and several other sub-companies for an adaptive program that overcomes the Y2K problem.
- **Parallelization of Gauss-Seidel Iterations:**
  Parallelization of Gauss-Seidel is a difficult problem in scientific computing. Proposed an algorithm that approximates to Gauss-Seidel iterations and makes them fully parallel. Later, designed and implemented a program using C, Fortran and EPX, on a parallel machine which uses those ideas.

REFERENCES

Prof. Ambuj K. Singh
Department of Computer Science
3119 Engineering I
University of California at Santa Barbara
CA 93106-5110
Email: ambuj@cs.ucsb.edu
Phone: (805)-893-3236

Prof. Amr El-Abbadi
Department of Computer Science
3119 Engineering I
University of California at Santa Barbara
CA 93106-5110
Email: amr@cs.ucsb.edu or
aabbadi@aucegypt.edu
Phone: (805)-893-4239

Prof. Divyakant Agrawal
Department of Computer Science
3119 Engineering I
University of California at Santa Barbara
CA 93106-5110
Email: agrawal@cs.ucsb.edu
Phone: (805)-893-4385

Prof. Ming Li
University of Waterloo
Department of Computer Science, DC2339
Waterloo, Ont. N2L 3G1
Phone: 1-519-888-4659
Email: mli@wh.math.uwaterloo.ca
Fax: 1-519-885-1208

Asst. Prof. Hakan Ferhatosmanoglu
Dept. of Computer and Information Science
The Ohio State University
395 Dreese Lab, Room # 689
2015 Neil Ave
Columbus, OH 43210, USA
Email: hakan@cis.ohio-state.edu
Phone: (614) 292-6377