Bilkent University
Computer Engineering Department

MSc and PhD Programs

Prof. Dr. H. Altay Güvenir
Research Areas

• Algorithms, artificial intelligence, big data, bioinformatics, cloud computing, computational biology, computational geometry, computer architecture, computer graphics, computer networks, computer vision, cryptography, data mining, data science, data security, database systems, graph visualization, high performance computing, image analysis, information retrieval, machine learning, mobile systems, parallel and distributed systems, pattern recognition, performance evaluation, scientific computing, virtual reality.
Applications

• Application Deadlines:
  Early: April 9, 2018; Regular: June 8, 2018

• Online Application:
  https://stars.bilkent.edu.tr/gradapp/

• Requirements for application:
  – CGPA ≥ 2.80 / 4.00
  – ALES (Turkish citizens) or GRE (Foreign applicants)
    • ALES: Quantitative ≥ 55 (for MS), 80 (for PhD w/o BS)
    • GRE: Quantitative ≥ 153 (MS), 157 (PhD)
  – English Proficiency: TOEFL (IBT) ≥ 87 or IELTS avg ≥ 6.5 (and min 5.5 in each section)
  – And YDS ≥ 55 (for Turkish citizen applying PhD).
Interview

- **Date:** April 16, 2018; June 18,19, 2018
- **Hard copy documents to be submitted during interview:**
  - Transcript
  - ALES or GRE score report
  - Proof of English Proficiency (TOFLE, IELTS or YDS report)
Acceptance requirements

- At least one faculty member willing to work with the applicant towards thesis
- Department approval
- Graduate School of Engineering and Science approval
Scholarship Options

- **Department scholarship**
  - Tuition waiver
  - Stipend
  - Accommodation support
  - Health Insurance
  - Additional Scholarship for Teaching Assistantship
  - Office (shared)

- **TÜBİTAK Scholarship or TÜBİTAK projects**
  - Tuition waiver
  - Accommodation
  - Health Insurance
  - Office (shared)
Scholarship Options

• Project grants (except TÜBİTAK projects)
  – Tuition waiver
  – Stipend (paid from the project budget)
  – Accommodation
  – Health Insurance (paid from the project budget)
  – Additional Scholarship for Teaching Assistantship
  – Office (shared)

• Graduate School scholarship
  – Tuition waiver
  – Additional Scholarship for Teaching Assistantship
Degree Requirements

• **MSc**
  - 7 technical elective courses + Seminar + MS Thesis + Research Methods and Academic Publication Ethics course

• **PhD**
  - 7 technical elective courses + Seminar + Qual Exam + PhD Thesis + Journal Publication Research Methods and Academic Publication Ethics course
Faculty Members

• In alphabetic order

(Please contact with them in person for details)
My current research is two pronged:

i. contextual reasoning in AI
ii. public implications of the Internet
Contextual reasoning in AI

Problems originating from natural language processing and understanding, especially those caused by the lack of explicit context:

When we say a particular thing, we do so in a context. Thus, there are embedded background assumptions available only through the context. We are also good at shifting between contexts. Can context be formalized in a formal framework? This would lead to improved software not only in Natural Language Understanding but also in knowledge-based systems, robotics, search engines, and personal assistants like Siri, Cortana, OK Google.

My grad courses CS 563 and CS 661 examine contexts, as well as other AI topics.
Public implications of the Internet

Problems caused by tensions between local versus global, individual versus corporate, and democratic versus authoritarian in the framework of the Internet. More specifically:

Basic concepts and history of the Internet as a socio-political medium. Free access to information in the networked world. The Internet as a trusted communications medium in light of security and privacy issues. Censorship and politics of the Internet vis-à-vis interventions by oppressive governments. Social movements in the 21st century marked by rapid broadcast of ideas and images.

My grad course CS 513 treats such societal aspects of the Internet.
Selim Aksoy
saksoy@cs.bilkent.edu.tr
http://www.cs.bilkent.edu.tr/~saksoy
Office: EA 423 (4th floor)

Research interests:
➢ Computer vision
➢ Pattern recognition
➢ Machine learning
➢ Data mining

Current topics:
➢ Remote sensing image analysis
➢ Image and video mining
➢ Medical image analysis

RETINA Vision and Learning Group
http://retina.cs.bilkent.edu.tr
Sponsored Research Projects

➢ Remote sensing image analysis
  ➢ TÜBİTAK CAREER Grant, 158,000 YTL, 2005-2010
  ➢ EC Joint Research Centre, 35,000 Euro, 2008
  ➢ FP6 Marie Curie Grant, 80,000 Euro, 2005-2007
  ➢ U.S. Army, $850,000, 2003-2005
  ➢ NASA Goddard Space Flight Center, $430,000, 2001-2004

➢ Image and video mining
  ➢ TÜBİTAK and COST 292 Action, 102,060 YTL, 2004-2008
  ➢ DPT, 2004-2005

➢ Medical image analysis
  ➢ TÜBİTAK CAREER Grant, 158,000 YTL, 2005-2010
  ➢ U.S. National Library of Medicine, $750,000, 2001-2004
Remote Sensing Image Analysis

- Increasing resolution in space and time (NASA receives 3 TB/day)
- Object recognition, classification and retrieval (Have you tried Google Earth?)

- Urban planning / monitoring
- Effects of natural disasters
- Environmental monitoring
Image and Video Mining

- Huge amounts of multimedia data (personal, commercial, free, …)
- Semantic classification and retrieval (Have you tried Google Video?)

- Object recognition
- Scene classification
- Combining image/audio/motion/text
Medical Image Analysis

- Microscopic image classification
- Cancer detection / grading
- Content-based searching of past cases
- Automatic report generation

Dr. Selim Aksoy
Combinatorial algorithms to analyze high throughput sequence data to discover, genotype, and phase genomic variants, assemble genomes and transcriptomes.

Test genome

Random shearing and Size-selection

Paired-end sequencing

Reference Genome

Resequencing

Read mapping and variation analysis

De novo sequencing

Contigs/Scaffolds

Assembly
Types of genomic variants

**SNP**: Single nucleotide polymorphism (substitutions)

**Indel**: Insertions and deletions of sequence of length 1 to 50 basepairs

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA CAG TGC GCG C - T</td>
<td>CA CAG TGC - G CAT</td>
</tr>
</tbody>
</table>

SNP deletion insertion

**Short tandem repeats**: microsatellites, minisatellites, alpha & beta satellites, etc.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>C A G C A G C A G C A G</td>
<td>C A G C A G C A G C A G</td>
</tr>
</tbody>
</table>

**Structural variation**: Genomic alterations > 50 bp
Deletions, insertions, mobile elements, duplications, inversions and translocations
Genome sequencers

- Roche/454
- AB SOLiD
- Illumina HiSeq2000
- Pacific Biosciences RS
- Illumina MiSeq
- Oxford Nanopore MinION
- Oxford Nanopore GridION
- Complete Genomics
- Ion Torrent PGM
- Ion Torrent Proton

... and more! All produce data with different properties.
Selected publications


**Combinatorial algorithms for structural variation detection in high throughput sequenced genomes.** Genome Research, Jul, 19(7):1270-8, 2009.


**A draft sequence of the Neandertal genome** Science, 7 May, 328 (5979):710-722, 2010.  
Recipient of the 2010 AAAS Newcomb Cleveland Prize.


**Limitations of next-generation genome sequence assembly.** Nature Methods, 8(1):61-65, 2011.  


Projects

- Discovery and characterization of genomic variation
  - Funded by EU Marie Curie Actions Career Integration Grant
- Algorithms and hardware designs for ultra-fast mapping of HTS reads to reference genome assemblies
  - Funded by US National Institutes of Health
- De novo and hybrid (multi-platform) sequence assembly.
- Genomic repeat discovery, classification and annotation.
- Distributed algorithms for genome assembly.

Positions available. Contact if you have B.Sc. or M.Sc. degree in computer science, computer engineering, electrical engineering, or mathematics, and if you are interested in combinatorial optimization, approximation algorithms, and graph theory. Strong programming skills in C/C++ are highly desired.

Successful applicants will also have a chance to contribute to many international consortiums such as the 1000 Genomes Project and the Genome 10K, and will involve in other international collaborations with researchers in Vancouver, Seattle, Barcelona, Bari, Pittsburgh, and more.

Basic understanding of biology/genetics/genomics is a plus, but not required.
Volume rendering of a combustion chamber: pressure and velocity fields.

Screen assignments for 24 processors: jagged and hypergraph partitionings.
Paralel and Grid Computing Group

Projects funded by TUBITAK, Intel and European Union

- EU-funded FP6 SEE-GRID2 project (2006-2009)
- Parallel Text Retrieval & Query Processing (TÜBİTAK, 2006-2008)
- EU-funded FP6 EUMedGrid project (2006-2008)
- EU-funded FP6 SEE-GRID project (SE4SEE application) (2004-2006)
- Task scheduling for PC clusters (TÜBİTAK, 1999-2002)
- Unstructured domain mapping (EU ITDC, 1995-1998)
- Parallel direct volume rendering (TÜBİTAK EEAG, 1995-1997)
Parallel and Grid Computing Group

- Parallel computing architectures
  - Intel iPSC/2 hypercube (32 nodes)
  - Parsytec CC24 (24 nodes)
  - PC cluster (Borg) (32 nodes)
  - PC cluster (Skynet) (48 nodes)
  - Grid cluster (16 nodes)
Paralel and Grid Computing Group
Current research topics

- Algorithm design for parallel and distributed computing applications
- Parallel text retrieval, query processing
- Parallel and distributed Web crawling
- High-performance geographical information systems
- Parallel scientific computing
- Inverted index compression
- Parallel and distributed data mining
- Task allocation and scheduling for Grid systems
- Grid-enabled Web search
- Parallel direct volume rendering

Paralel and Grid Computing Group

- Contact information
  - Cevdet Aykanat
  - Phone: 312 290 1625
  - Email: aykanat@.cs.bilkent.edu.tr
  - Homepage: http://www.cs.bilkent.edu/~aykanat
Bilkent Information Retrieval Group

Faculty
Fazlı Can
Seyit Koçberber

Graduate Students
Cem Aksoy
Ceyhun Karbeyaz
Çağrı Toraman
Anıl Türel
Ahmet Yeniçağ

Undergraduate Students
Turgut Işık
Oğuz Kaya
Harun Özden
Abdullah Şahin

Volunteers
Aykut Bal
Akif Buğday
Bilge Köroğlu
Hasan Can Tuncay

Some Prev. Members
Erman Balçık
Ahmet Buğdayıcı
Tunay Gür
Cihan Kaynak
Levent Koç
İbrahim Uysal

Other Contributors
Cevdet Aykanat
Pınar Duygulu
Özgür Ulusoy
İsmail Şengör Altingövde
Özgür Bağlıoğlu
Ethem F. Can
Gönenç Ercan
Süleyman Kardaş
H. Çağdas Öcalan
Erkan Uyar
Research Interests

- New Event Detection and Tracking
- Novelty Detection
- Information Retrieval
- Information Filtering
- News Categorization
- Text Mining & Processing
Bilkent Haber Portalı

Ana Sayfa | Ürünler | Yardım | Hakkımızda

KATEGORİLER

Ekonomi
Politika
Türkiye
Dünya
Spor
Kültür - Sanat
Sağlık
Bilim Teknoloji
Yazarlar

SON HABERLER

KÜÇÜK TAML ASKERLER SINIFLARINA DÖNÜYOR...
Sri Lanka'da 30 yıllık çatışmaların ardından hükümet tarafından eski topraklarına döndürülen ülkenin doğu ve kuzey [Devamı...]

KÜTAHYA'DA GÖÇÜK NEYDANA GELEN MADEN...
Kütahya'nın Tavşanlı ilçesinde, göçük neydana gelen madden ocağı hakkında Valilikçe yaklaşık 2,5 ay önce eksikleri [Devamı...]

KUROYEMİŞİN AZI KARAR, ÇOCUZ ZARAR...
İkili tıbbında dünya tıncısı olan ugunsuz kuru yemiş, öğütlü miktarda alınan sağlıklık yaşamın katkı sağlayacağı bildirildi. [Devamı...]

VALİ GÜLER: DEYLEİN VERDİĞİ GÖREVİN...

GÜNCELI & GEÇMİŞ OYAYIN

Güçlü Olaylar

- KEİTA MILLİ TAKİMDA... İzleyenler (54)
- YENİ BİR SERİN KATIL... İzleyenler (7)
- AĞLAYAN ÖĞRENCİ VALİYE ZOR... İzleyenler (6)
- GÜNVEDE PLAJLAR DÖLDÜ TAŞTLI... İzleyenler (7)
- ÇÖRÜM'DA KENEDEN I ÖLÜM... İzleyenler (9)
- ŞOK İDDİYA: MICHAEL JACKSON... İzleyenler (9)
- PAPAKOSTANDINIU: YENİ ÖNELMELER ALINMASI... İzleyenler (7)
- 10 PROJENİN ŞAMPİYONU VARYAP... İzleyenler (5)
- ARNAVUTLUK'TA SIYASYI KRİZ... İzleyenler (6)
- TÜRKİYE VE YUNANISTAN, ORTAK... İzleyenler (5)
- KARAPARADA İNCELEME HAVUZU'NDAYIZ... İzleyenler (7)
- KÖNÜT KREDİSİ FAZINDE YÜKSELIŞ... İzleyenler (8)
- HİDRO DÜŞMAN KARDEŞLER MALLARI... İzleyenler (15)

EN ÇOK OKUNANLAR
You are the right person for our group

- If you are good at four core practices of CS
  - programming,
  - systems thinking,
  - modeling, and
  - innovation.

- If you
  - "can dream,"
  - "can do," and
  - "can write."
A. Ercument Cicek

cicek@cs.bilkent.edu.tr

cs.bilkent.edu.tr/~cicek

My research focuses on developing algorithms for analyzing biological data. Using various biochemical networks, I analyze the data and make discoveries in molecular biology. The search space is large for even a basic machine learning task. I design algorithms that can discover biomarkers in particular for genetic diseases. The data is vast and complex, requiring efficient analysis. Central to this is the lack of computational power in the bio-research data. Today, the bottleneck in the 'big data' is the informatics. Consequeently, the cost of bioscience is huge. For example, the first genome took 13 years and $1b to sequence. Even though various technologies have resulted, it takes a day to sequence the genome. My work is relevant now in accumulating data and making it tractable.
A. Ercument Cicek

cicek@cs.bilkent.edu.tr

Understand the genetic architecture.

Algorithm needed to discover new genes to biological networks along with intelligent search.

So far we have only discovered ~50 of them.

Gene Discovery for Autism Spectrum Disorder
A. Ercument Cicek

cicek@cs.bilkent.edu.tr
cs.bilkent.edu.tr/~cicek

Using Dynamic Network Algorithms to Model Neurodevelopment

-Autism is a neurodevelopmental disorder that affects the brain's development, particularly during the fetal period.

-We design algorithms to model the evolving gene interaction networks and understand the complexity of autism's dynamics.

-Our algorithms analyze the early childhood networks to gain insights into the functional aspects of autism.
Metabolic Networks to Understand Cancer

Metabolites are the small compounds in the body and have been found to be key biomarkers to define certain tumors. We use network algorithms and build online systems that analyze metabolic signatures in tumors and understand the differences within the subtypes of the same cancer.
A. Ercument Cicek
cicek@cs.bilkent.edu.tr
cs.bilkent.edu.tr/~cicek

---

Selected Publications:

- **De Novo Chip-Seq Analysis**, Biology 2015, 16:205.
- **Expected metabolite to determine molecular autism**, PloS Computational Biology 2013, 9(1):
  e1002859.
- **Molecular autism expected level changes using metabolite,,** PloS Computational Biology 2013, 9(1):
  e1002859.
Tuğrul Dayar
(tugrul@cs.bilkent.edu.tr)

- Performance modeling and analysis
- Scientific computing (especially numerical linear algebra for stochastic matrices)
- Bioinformatics
- Computer networks

requires solid background:
1) in computer engineering
2) specifically in probability theory, linear algebra, numerical analysis, and high-level programming,
which one must either have or be willing to develop

Visit: www.cs.bilkent.edu.tr/~tugrul/tugrul.html for further information
Projects:  

• **PATIKA & Visibio** [Pathway Modeling, Analysis & Visualization Tools]  
• **Chisio** [Graph Visualization Tool]

Our projects have been supported by TÜBİTAK and Tom Sawyer Software (CA, USA)
Social Network of Drug Traffickers
The PATIKA Project
www.patika.org

Sample PATIKA Tool: PATIKAweb

Lost in the jungle of cellular processes...?

PATIKAweb can show you the PATHway
Interested in being part of this kind of research & development? Then join us! Please contact me at ugur@cs.bilkent.edu.tr
Data Management and Mining
Hakan Ferhatosmanoglu

- We investigate data management and data mining methods for emerging applications:
  - Data streams (telecom data warehouses, financial markets)
  - Social Media (Twitter text and social network)
  - Bioinformatics (proteins, genes)

- We aim to build scalable systems for online exploration of text, image, time-series

- Funded by USA Dept of Energy, NSF, NASA, Pfizer, IBM
Data Management and Mining
Hakan Ferhatosmanoglu

• Looking for students with theoretical interest and systems building skills (basically math and programming..)

Please contact me for more information: hakan@cs.bilkent.edu.tr
We are conducting research in data-intensive distributed systems, looking at issues such as parallelization, load balancing, load shedding, placement optimization, fault-tolerance, etc.

Particular topics of interest include:

- Distributed data stream processing systems
- Big Data technologies, such as distributed key-value stores, map/reduce systems, bulk-synchronous parallel processing frameworks, etc.
- Large-scale distributed graph management and mining
- Peer-to-peer, mobile, and sensor data management
Data Stream Processing Systems

Stock market
- Impact of weather on securities prices
- Analyze market data at ultra-low latencies

Law Enforcement, Defense & Cyber Security
- Real-time multimodal surveillance
- Situational awareness
- Cyber security detection

Fraud prevention
- Detecting multi-party fraud
- Real time fraud prevention

Transportation
- Intelligent traffic management

Telephony
- CDR processing
- Social analysis
- Churn prediction
- Geomapping

Health & Life Sciences
- Neonatal ICU monitoring
- Epidemic early warning system
- Remote healthcare monitoring

Manufacturing
- Process control for microchip fabrication

Natural Systems
- Wildfire management
- Water management

Telephony
- CDR processing
- Social analysis
- Churn prediction
- Geomapping

Other
- Smart Grid
- Text Analysis
- Who’s Talking to Whom?
- ERP for Commodities
- FPGA Acceleration

e-Science
- Space weather prediction
- Detection of transient events
- Synchrotron atomic research
Automatic Parallelization

- Transparent parallelization: Locate parallel regions in streaming applications without user intervention
- Safe parallelization: Generate parallel alternatives that produce the exact same results as the sequential version
- Elastic parallelization: Dynamically adjust profitability decisions w.r.t. parallelization based on run-time dynamics
Looking for students

We are seeking highly motivated and skilled students for graduate level research in data intensive distributed systems. Both Ph.D. and Masters level students are welcome. Students with strong interest in systems and implementation as well as a good understanding of fundamental theory are encouraged to apply.

Contact: www.bugragedik.com
Selected Publications


Main research interests:

- Medical image analysis for automatic cancer diagnosis, grading, and prognosis
- Machine learning for intelligent medical systems
Current projects:

- **Diagnosis and grading of colon cancer**
  - Construction of new biocomputational methods
  - Colon cancer diagnosis
  - Colon cancer grading
  - Tissue image segmentation
  - Colon gland segmentation

- **Designing medical diagnostic systems**
  - Cost-sensitive classification
  - Qualitative decision theory
  - Dynamic model selection and combination
Research Topics:
1. Augmented and Virtual Reality
2. Terrain and Urban Modeling and Visualization
3. Human Modeling and Animation
Augmented and Virtual Reality

- Crowd Simulation
- Realistic Lighting
- Camera Registration and Tracking
- Augmented Reality on Mobile Devices
Terrain and Urban Modeling and Visualization

- Level-of-detail management
- View-dependent refinement
- Stereoscopic visualization
- GPU-based tessellation
- Crowd simulation in urban environments
Human Modeling and Animation

Real-time Virtual Garment Fitting Using Depth Sensor Data

Realistic Rendering of Joint Regions Using Bone Splitting

A Model Wearing a Sundress with Different Postures

A Model Wearing a Vest and Jeans with Different Postures
Current research topics include

Learning to

- rank instances
- model risk factors
- estimate risks
- suggest to increase success

Application areas: Medical, Social Networks
Faculty Member
Ibrahim Korpeoglu
Assoc. Professor
Dept of Computer Engineering
Bilkent University

Email: korpe@cs.bilkent.edu.tr
Web: http://www.cs.bilkent.edu.tr
Office: Engineering EA 409
Phone: 290 25 99

Working on Problems and Projects related with Computer Networks and Computer Systems

Research Areas:
- Computer Networks
- Computer Systems
- Network and Distributed Algorithms
- Wireless Networks
- Distributed Systems
- Cloud Computing
- P2P Networks
- Sensor Networks
- Internet of Things
- WiFi and Bluetooth
- Big Data Systems
Efficient **Resource Allocation in Heterogeneous Cloud Infrastructures**
Sponsor: TUBITAK

Supporting Real-time Traffic in **Wireless Ad Hoc and Sensor Networks**
Sponsor: TUBITAK

**Bluetooth Scatternet Construction and Bluetooth Applications**
Sponsor: TUBITAK

Network Middleware for Environmental Monitoring and Control with **Wireless Ad hoc, Mesh and Sensor Networks**
Sponsor: IBM Corporation

**Intel WCNC, Wireless Networking Curriculum Enhancement Project**
Sponsor: Intel Corporation

**FIRESENSE Fire Detection and Management through a Multi-Sensor Network** for the Protection of Cultural Heritage Areas from the Risk of Fire and Extreme Weather Conditions
Sponsor: European Commission FP7 Programme, ENV

Network of Excellence in **Wireless Communications (NEWCOM and NEWCOM++)**
Sponsor: European Commission FP7 Programme, ICT
Wireless Mesh Networks
- Routing
- Channel assignment
- Interference modeling
- Interference mitigation

Sensor Networks
- Energy efficient routing
- Activity scheduling
- Channel access scheduling
- ZigBee wireless technology
- ZigBee routing

P2P Networks
- Query forwarding
- Free riding
- File sharing and lookup

Delay tolerant networks
- Routing and Scheduling

Testbeds

Cloud Computing
- Resource allocation
- VM placement
- Network virtualization
- Network embedding
Hidayet Aksu, Mustafa Canim, Yuan-chi Chang, Ibrahim Korpeoglu, Ozgur Ulusoy, *Distributed k-Core View Materialization and Maintenance for Large Dynamic Graphs*, IEEE Transactions on Knowledge and Data Engineering, 2014.


High-Performance and Energy Efficient Computing

*Algorithms, Systems, and Applications*
Active projects:

- Energy efficient FPGA accelerators for big data applications \textit{(supported by European Commission, in collaboration with Intel, Oregon)}
- Parallelizing Collaborative Filtering algorithms for recommender systems
- Parallel and vectorized scientific computing algorithms
- Optimizing memory architecture for graph analytics applications for large datasets.

I am interested in graduate student candidates who have one or more of the following:

- Strong algorithmic background and analytical skills
- Interest in solving programming puzzles
- C++ programming skills
- Experience with multi-core, GPU, and/or FPGA programming

For more information, see \texttt{www.bilkent.edu.tr/~mustafa.ozdal}. You can send an email to \texttt{mustafa.ozdal@cs.bilkent.edu.tr} to set up an appointment.
Özcan Öztürk
Office: EA 421 Phone: 290-3444
Email: ozturk@cs.bilkent.edu.tr
URL: http://www.cs.bilkent.edu.tr/~ozturk

- **Computer architecture** - memory scheduling, memory hierarchy design, metrics to consider energy, performance, reliability....
- **Multicore/Manycore architecture** - design of multicore systems, application mapping, data mapping, communication
- **Heterogeneous computing** - heterogeneous multicore design, core and cache selection, application execution
- **Parallel programming/systems/applications** - OpenMP, MPI, GPGPU, application characterization, automatic parallelization, scheduling
- **Cloud computing** - at the system level and architectural optimizations, heterogeneity aware scheduling
- **Embedded computing** - energy, performance, ...
- **Compiler optimizations** - code modifications and optimizations to generate better applications
Current Projects

- Heterogeneous Multicore Design  
  *Funding: EC FP7*

- Parallelization for Heterogeneous Multicore Architectures  
  *Funding: IBM*

- Utilizing Accelerator Technologies in the Cloud  
  *Funding: Türk Telekom*

- Parallelizing Data Mining applications using GPUs  
  *Funding: Nvidia*

- Accelerator Design for Graph Parallel Applications  
  *Funding: Intel*
Database Research
Özgür Ulusoy
http://www.cs.bilkent.edu.tr/~oulusoy/

- Web Databases and Search Engines
- Multimedia Databases
- Big Data and Social Network Analysis
Web Databases and Search Engines

- Social Web search and personalization
- Domain-specific search engines
- Efficiency and scalability issues for Web Search Engines (caching, index pruning)
- Web information extraction
- Modeling and querying of Web resources
- XML querying & searching

http://www.cs.bilkent.edu.tr/~bilweb
Search Engines are the key to access Web Data

S. Brin meets L. Page
1995

First billion-URL index
The world’s largest!
≈5000 PCs in clusters!

2000

Google counts 1 trillion unique URLs
2008

First billion-URL index
The world’s largest!
≈5000 PCs in clusters!

1995

Birth of Google
1998

Index grows to 4.2 billion pages
2004

TBs or PBs of data/index
Tens of thousands of PCs
2009

The size of the indexed World Wide Web
(Number of webpages)


Efficient and scalable strategies are of vital importance!
Multimedia Databases
(joint work with Prof. Uğur Gündükbay)

- Video Retrieval Systems
- Mobile Visual Search
- Ottoman Archive Content-Based Retrieval System

http://www.cs.bilkent.edu.tr/~bilmdg
**BilVideo**: Integrated video DBMS supporting low-level, spatio-temporal, motion and semantic querying of videos

**System Architecture**

- **Fact-Extractor**
  - Extracted Facts
  - **Knowledge-Base**
  - **Query Processor**
    - Query
    - Results
  - **Feature Database**
  - **Video-Annotation**
    - Videos
    - **Raw Video Database (File System)**
  - **Object-Relational DBMS**

**Query GUIs**

- **SPATIO-TEMPORAL QUERY GUI**
- **SEMANTIC QUERY GUI**
- **TRAJECTORY QUERY GUI**
BilVideo-7: An MPEG-7 Compatible Video Retrieval System

Users

Web Client

Visual Query Interface

Raw Video Database (File System)

Video Processing
- SBD, Segmentation
- Object Extraction
- Annotation
- etc.

Feature Extraction
- MPEG-7 Features

Query Processor

XML-Native Feature Database (Tamino)

Automatic processing: segmentation, tracking, feature extraction, annotation, indexing

Powerful querying capability for video data
- keyword and content-based queries
- spatio-temporal object queries

Example query formulation
- Salient video object extraction

keywords: trees, greenery, sky – bush, putin, dog
Mobile Image Search Using Multi-Image Queries

Workflow of the Search System

Early and Late fusion methods

Multi-View Dataset and Queries

Single (a) and multi-view queries and corresponding result lists using early (b) and late (c) fusion methods.
Big Data and Social Network Analysis

- **Social Network Data Analysis on Big Data Processing Platforms**
  - Development, implementation and evaluation of algorithms/methods to process/analyze social network data for various social network problems. (joint work with Prof. İbrahim Körpeoğlu)

- **Decentralized Social Networks**
  - Development, implementation and evaluation of algorithms/methods for peer-to-peer social networks. (joint work with Prof. Hakan Ferhatosmanoğlu)