

COLLEGE OF ENGINEERING

Annual Activities Report

1 NAME:

Alin V. Dobra

2 TENURE:

Tenured

3 CURRENT RANK:

Associate Professor, Computer and Information Science and Engineering

4 DATE APPOINTED TO THIS RANK:

August 2009

5 AFFILIATE APPOINTMENTS:

None

6 ASSIGNED ACTIVITY

	Spring 2013	Summer 2013	Fall 2013
Teaching	45%	2%	50%
Research	40%	90%	35%
Service	15%	8%	15%
Total	100%	100%	100%

7 TEACHING, ADVISING AND/OR INSTRUCTIONAL ACCOMPLISHMENTS

Teaching Philosophy: My main strength as a teacher, and a real asset as a researcher, is an broad understanding of Computer Science as a whole rather than just a few subjects, as it is often the case. This breadth of knowledge allows me to easily teach a variety of subjects well outside my research area.

- **COP 5615 Distributed Operating System Principles** This is a graduate level course focused on distributed systems and modern applications like large ecommerce sites, peer-to-peer systems, and distributed computation. My overall instruction evaluation for this class was 4.49 out of 5.0
- **COP 6726 Database System Implementation** This is a very advanced class on database system implementation. Students are required to implement a full fledged database that can execute SQL and were taught classic and modern takes on how to implement high performance databases. My overall instruction evaluation for this class was 4.48 out of 5.0

Supervising: I supervised 3 Ph. D students during the last 1 year. The active students are Supriya Nirkhiwale and Andrei Todor. Andrei is co-advised with Tamer Kahveci. Louis Cheung, my 3rd student decided to abandon the Ph-D program.

8 TEACHING EVALUATIONS

Course: COP 5615
Fall, 2013
High=5, Low=1

Required: NO
Enrolled: 152
Responses: 37

Question	Dr. Dobra Mean	Dept. Mean	College Mean
1. Description of course objectives and assignments	4.41	4.12	4.24
2. Communication of ideas and information	4.35	3.97	4.05
3. Expression of expectations for performance in this class	4.49	4.07	4.17
4. Availability to assist students in or out of class	4.43	4.18	4.19
5. Respect and concern for students	4.49	4.29	4.32
6. Stimulation of interest in course	4.57	4.04	4.08
7. Facilitation of learning	4.49	4.00	4.03
8. Enthusiasm for the subject	4.73	4.32	4.38
9. Encouragement of independent, creative, and critical thinking	4.76	4.19	4.18
10. Instructor Overall	4.49	4.09	4.17

Course: COP 6726
Fall, 2012
High=5, Low=1

Required: NO
Enrolled: 112
Responses: 48

Question	Dr. Dobra Mean	Dept. Mean	College Mean
1. Description of course objectives and assignments	4.29	4.06	4.15
2. Communication of ideas and information	4.46	3.94	4.01
3. Expression of expectations for performance in this class	4.42	4.04	4.13
4. Availability to assist students in or out of class	4.31	4.13	4.14
5. Respect and concern for students	4.52	4.22	4.26
6. Stimulation of interest in course	4.68	4.01	4.07
7. Facilitation of learning	4.40	3.96	4.03
8. Enthusiasm for the subject	4.70	4.31	4.37
9. Encouragement of independent, creative, and critical thinking	4.69	4.18	4.20
10. Instructor Overall	4.48	4.05	4.15

9 GRADUATE COMMITTEE ACTIVITIES

So far, I have served in 70 committees (9 as Chair, 3 as Co-Chair, 1 External Member, 55 as Member and 2 as Minor).

Role	Student Name	Degree	Major	Degree Date
Chair	ALEX,DALEY	M.S.	Computer Engineering	8/14/2012
Chair	CHEN,LIXIA	Ph.D.	Computer Engineering	12/20/2011
Chair	CHEUNG,LOUIS	Ph.D.	Computer Engineering	
Chair	DHURANDHAR,AMIT S	Ph.D.	Computer Engineering	8/11/2009
Chair	GOLANI,GURUDITTA	M.S.	Computer Sciences	8/6/2005
Chair	JAMPANI,RAVINDRANATH	Ph.D.	Computer Engineering	8/14/2012
Chair	KIM,EUNKEE	Ph.D.	Computer Engineering	
Chair	NIRKHIWALE,SUPRIYA	Ph.D.	Computer Engineering	
Chair	RUSU,FLORIN I	Ph.D.	Computer Engineering	5/5/2009
Chair	TODOR,ANDREI	Ph.D.	Computer Engineering	
Chair	ZHENG,YONGJIE	Ph.D.	Computer Engineering	
Co-Chair	CHITNIS,LAUKIK VILAS	Ph.D.	Computer Engineering	8/12/2008
Co-Chair	MARTINEZ,ALEXANDRA MARIA	Ph.D.	Computer Engineering	12/18/2007
External	YAO,KAI	Ph.D.	Materials Sc. and Eng.	
Member	ABBASIMOGHADDAM,SAEED	Ph.D.	Computer Engineering	12/18/2012
Member	ARUMUGAM,SUBRAMANIAN	Ph.D.	Computer Engineering	8/12/2008
Member	AY,FERHAT	Ph.D.	Computer Engineering	8/9/2011
Member	BAKER,TROY A	Ph.D.	Computer Engineering	
Member	BANDYOPADHYAY,NIRMALYA	Ph.D.	Computer Engineering	12/20/2011
Member	CHEN,TAO	Ph.D.	Computer Engineering	8/14/2012
Member	CHEN,YANG	Ph.D.	Computer Engineering	
Member	CHOU,HONGCHI	Ph.D.	Computer Engineering	
Member	EOM,BOYUN	M.S.	Computer Sciences	8/6/2005
Member	GABR,HAITHAM MOHAMMAD	Ph.D.	Computer Engineering	
Member	GRANT,CHRISTAN EARL	Ph.D.	Computer Engineering	
Member	GULSOY,GUNHAN	Ph.D.	Computer Engineering	8/13/2013
Member	HAN,SEUNG CHUL	Ph.D.	Computer Engineering	5/5/2007
Member	HSU,WEI-JEN	Ph.D.	Computer Engineering	8/12/2008
Member	JOSEPH,REJITH G	M.S.	Computer Engineering	5/3/2011
Member	JOSHI,SHANTANU SHARAD	Ph.D.	Computer Engineering	8/14/2007
Member	KANJILAL,VIRUPAKSHA	Ph.D.	Computer Engineering	12/18/2012
Member	KHAN,MD ARIFUL	M.S.	Computer Engineering	8/10/2010
Member	KHANAPURE,VISHAL A	M.S.	Computer Engineering	8/11/2009
Member	KIM,EUNJU	Ph.D.	Computer Engineering	5/7/2013
Member	KUMAR,UDAYAN	Ph.D.	Computer Engineering	12/18/2012
Member	LI,BO	Ph.D.	Computer Engineering	8/14/2012
Member	LI,KUN	Ph.D.	Computer Engineering	
Member	LI,XUEHUI	Ph.D.	Computer Engineering	12/18/2007
Member	LIU,HECHEN	Ph.D.	Computer Engineering	12/18/2012
Member	LIU,JUN	Ph.D.	Computer Engineering	5/6/2008
Member	LU,I-HSUAN	Ph.D.	Computer Engineering	
Member	MCKENNEY,MARK A	Ph.D.	Computer Engineering	8/12/2008
Member	MOOLA,ANIL	M.S.	Computer Engineering	5/6/2008
Member	MUN,MIN YOUNG	Ph.D.	Computer Engineering	
Member	PANSARE,NIKETAN R	M.S.	Computer Engineering	12/22/2009
Member	PAULY,ALEJANDRO	Ph.D.	Computer Engineering	5/5/2007
Member	PEREZ,LUIS L	Ph.D.	Computer Engineering	
Member	PEREZ,LUIS L	M.S.	Computer Engineering	12/22/2009
Member	POL,ABHIJIT A	Ph.D.	Computer Engineering	8/14/2007
Member	PRAING,REASEY	Ph.D.	Computer Engineering	8/12/2008

Role	Student Name	Degree	Major	Degree Date
Member	RAJAMANICKAM,SIVASANK	Ph.D.	Computer Engineering	12/22/2009
Member	RASHEED,HASSAN S	Ph.D.	Computer Engineering	5/5/2009
Member	RAVUNNIKUTTY,GIRISH	M.S.	Computer Engineering	5/3/2011
Member	SEN,SOMAK	M.S.	Computer Engineering	5/5/2009
Member	SHRIVASTAVA,KARTIK P	M.S.	Computer Engineering	8/10/2010
Member	SOMAIYA,MANAS H	Ph.D.	Computer Engineering	12/22/2009
Member	THAKUR,GAUTAM	Ph.D.	Computer Engineering	12/18/2012
Member	VENKATESWARAN,JAYENDRA G	Ph.D.	Computer Engineering	12/18/2007
Member	VISWANATHAN,GANESH	Ph.D.	Computer Engineering	12/20/2011
Member	WANG,YIBIN	Ph.D.	Computer Engineering	
Member	WEINRICH,BRIAN ERWIN	Ph.D.	Computer Engineering	5/5/2007
Member	WU,MINGXI	Ph.D.	Computer Engineering	8/12/2008
Member	XU,FEI	Ph.D.	Computer Engineering	8/11/2009
Member	YERLAN,SENCER NURI	Ph.D.	Computer Engineering	
Member	YUAN,WENJIE	Ph.D.	Computer Engineering	8/9/2011
Member	YUN,YOUNGSANG	Ph.D.	Computer Engineering	8/10/2010
Member	ZANDI,HELIA	Ph.D.	Computer Engineering	
Member	ZENG,QI	Ph.D.	Computer Engineering	
Member	ZHANG,XU	Ph.D.	Computer Engineering	12/18/2007
Minor	VEERAMANI,KARTHIK	M.S.	Electrical and Computer Eng.	12/16/2006
Minor	VELUCHAMY,NIVETHA	M.S.	Electrical and Computer Eng.	5/4/2010

10 CONTRIBUTION TO DISCIPLINE/RESEARCH NARRATIVE

With the advent of cheap storage and fast networks, data is produced and stored at high speeds. The next big challenge in computer science is to process efficiently these large volumes of data. My research goal is to develop mathematical methods that allow design and analysis of algorithms that can process such large volumes of data, on one hand, and deeper understanding of existing algorithms on the other hand. Specific contributions that are the subject of my work are:

- Processing of data arriving at high speeds in large volumes, for example for data generated by computer networks, depends crucially on the ability to accurately summarize the data *on the fly*. I am developing techniques for producing provably good summarization and mathematical infrastructure that speeds up the design of such techniques. My work in this area is supported by an NSF CAREER award.
- Large databases are prevalent nowadays – Walmart records and stores every customer and warehouse transaction. Processing queries for such large databases is problematic since techniques from traditional database research lead to large running times for the queries. An alternative is to process the queries on a sample from the database and to estimate the result – while this idea is simple in principle, it leads to complicated analysis and a complete redesign of query processing engines. My work in this direction, together with Dr. Chris Jermaine, was published in SIGMOD 2005,06,07,08 and VLDB 2005 conferences. It is also supported by an NSF grant for the next 4 years.

An interesting new directions this work lead to is the development of very fast databases. The result with Chris Jermaine, published in SIGMOD 2010, rivals existing database solutions on data in the 10 TB range for a fraction of the cost – a 40,000\$ computer running DataPath rivals a 10 million \$ IBM DB2 solution for certain analytical queries.

Most of my work is published in highly competitive conferences (with low acceptance rate) and journals. A number of my papers had a significant impact in the community, as reflected by the number of citations

they received:

- *Gossip-based computation of aggregate information*: 937 citations
- *Processing complex aggregate queries over data streams*: 330 citations
- *SECRET: a scalable linear regression tree algorithm*: 72 citations
- *Sketch-based multi-query processing over data streams*: 46 citations

I have 1824 citations overall, an h-index of 14 and i10-index of 21.
(<http://scholar.google.com/citations?user=smE00BsAAAAJ>)

11 CREATIVE WORKS OR ACTIVITIES

I developed/help develop three major software systems:

- BioVerto: a tool for visualizing and analyzing biological networks (<http://bioverto.org>). January 2013 (joint effort with Tamer Kahveci)
- DBO: a database system that estimates the query result while the query is processed. This system is developed in collaboration with Chris Jermaine. 2006-2013
- DataPath: a fast database system particularly suited for large analytical query loads. This system is developed in collaboration with Chris Jermaine. 2009-now
- GLADE: an addon on top of DataPath that allows sophisticated data processing as part of a relational database engine. Collaboration with Sanjay Ranka, 2011-now

12 PATENTS AND COPYRIGHTS

1. *Sketch-based multi-query processing over data streams*, Alin Dobra, Johannes Gehrke, Rajeev Rastogi, Minos Garofalakis, Patent number: 7328220, Issue date: Feb 5, 2008

This patent covers the innovations in terms of extending a particular approximation technique, sketching, to multiple queries that are run simultaneously. The main innovation is a summary sharing technique that is useful for a wide class of approximations.

13 PUBLICATIONS

- a. **Books, Sole Author** - N/A
- b. **Books, Co-authored** - N/A
- c. **Books, Edited** - N/A
- d. **Books, Contributor**
- e. **Monographs** - N/A
- f. **Refereed Publications**

Journal Publications

1. Andrei Todor, Haitham Gabr, Alin Dobra, Tamer Kahveci. Large scale analysis of signal reachability. *Bioinformatics*. to appear

2. Andrei Todor, Alin Dobra, Tamer Kahveci. Characterizing topology of probabilistic biological networks. IEEE/ACM Transactions on Computational Biology and Bioinformatics Journal (IEEE/ACM TCBB). 10:4. 2013
3. Andrei Todor, Alin Dobra, Tamer Kahveci. Probabilistic Biological Network Alignment. IEEE/ACM Transactions on Computational Biology and Bioinformatics Journal (IEEE/ACM TCBB). 10:1, 2013
4. Supriya Nirkhivale, Alin Dobra, Christopher M. Jermaine: A Sampling Algebra for Aggregate Estimation. PVLDB 6(14): 1798-1809 (2013)

Refereed Proceedings

1. Haitham Gabr, Andrei Todor, Helia Zandi, Alin Dobra, Tamer Kahveci. PReach: Reachability in Probabilistic Signaling Networks. International Conference on Bioinformatics and Computational Biology (ACM-BCB), 2013

14 LECTURES, SPEECHES OR POSTERS PRESENTED AT PROFESSIONAL CONFERENCES/MEETINGS

International

N/A

State

N/A

Local

N/A

15 CONTRACTS AND GRANTS SINCE LAST PROMOTION OR DURING THE LAST FIVE YEARS

Funded Research Contracts and Grants

a. Funded Externally

Title of Grant	Funding Agen.	PI	Start-End	Value	Funding Portion
IIS: EAGER: A Framework for Large Data Analysis	NSF	PI	09/01/11-/8/31/14/	\$100,000	\$100,000
CIF: EAGER: Modeling Prob Bio Networks	NDF	Co-PI	08/13-07/15	\$175,000	\$87,500

Summary of Grant Funding Received, Mar 2013-present

Role	Total
Principal Investigator	\$100,000
Co-Principal Investigator	\$87,500
Totals	\$187,500

- b. **Funded Internally** N/A
- c. **Submitted, Pending Decision**
N/A
- d. **Submitted Not Funded**

Title of Grant	Fund. Agen.	PI	Start-End	Value Funding
BIGDATA: Small: DA: GLADE: A Framework for..	NSF	Co-PI	01/01/13-12/31/15	\$750,000
II-NEW: Research infrastructure for Cost Effective..	NSF	Co-PI	04/01/13-03/31/15	\$600,022

16 UNIVERSITY GOVERNANCE AND SERVICE

16.1 Department Committee Memberships

- Colloquium Committee
- Chair of Scholarships and Awards Committee

17 CONSULTATIONS OUTSIDE THE UNIVERSITY

N/A

18 EDITOR OF A SCHOLARLY JOURNAL, SERVICE ON AN EDITORIAL ADVISORY BOARD OR REVIEWER FOR A SCHOLARLY JOURNAL

Reviewer for the following journals:

- IEEE Transactions on Parallel and Distributed Computing

19 INTERNATIONAL ACTIVITIES

N/A

20 SERVICE TO SCHOOLS

N/A

21 MEMBERSHIP AND ACTIVITIES IN THE PROFESSION

- *Program committee* member for the following conferences:
 - International Conference on Data Engineering, 2014
 - ACM-SIGMOD 2014

22 HONORS

23 THE FURTHER INFORMATION SECTION

Signature_____

Date_____