1/4/05
******(note: updated versions will be placed on the web!)******

DARTMOUTH COLLEGE
COMPUTER SCIENCE DEPARTMENT
Winter, 2005

Current Trends and Ethical Issues in Computer Science (CS99)

Faculty: Fillia Makedon
Sudikoff 109
(603) 646-3048
Fillia.Makedon@dartmouth.edu
Office Hours: Monday 2-3 pm, Wednesday 11-12 noon or by appointment.

Teaching Assistant:
Yurong Xu
Thursday 4-5 pm, Friday 11-12 am or by appointment

Overview

This course is the culminating experience that caps your undergraduate computer science major and prepares you for an ethical career in computer science - a career that is not divorced from ethical and social responsibilities. Given the central role that computer science is playing in every aspect of our daily lives, it is important that you, as a graduating computer scientist and future leader in the field, be prepared to solve not only technical problems, but also problems that arise from the technology. The latter are far more complex and require practiced critical thinking skills and a deeper understanding of the issues. Thus, a good computer scientist must also be able to understand and analyze the impact of the computer technology he/she uses and produces and must be able to do original research that may include programming to test a hypothesis, critical analysis, modeling of the data, writing about the problem and evaluating the results of his/her research as well as defending the actions related to a particular technology.

Class Meetings

Sudikoff 115: Tuesdays and Thursdays at 2:00 – 3:50 PM
x-hour, 4:15-5:20 PM, Scheduled, unless otherwise noted

Course Goals

- To make you aware of ethical considerations and computer trends and to encourage commitment to ethical behavior ;
- To enhance your communication skills, both oral and written, including the presentation of coherent and consistent arguments;

- To make you capable of critical thinking and train you in the analysis of the materials covered. This includes raising your awareness of historical and current trends, improving your ability to analyze technology trends, helping you learn to sift out the ethical issues in a situation, to identify moral implications, and to clarify the different choices; and
- To develop your ability to evaluate the design and implementation of computer systems in a moral, social or legal context and to foresee problems that may arise as best as possible.
- To develop your ability to design an experiment that tests or models the impact of a technology and then write a report on the work that includes motivation, background research, new work, evaluation, future work and conclusions.

Expectations

You are seniors and expectations are high. You are expected to attend all classes, to be on time for class, and to be prepared for class. You are expected to manage the time for your assignments and be prepared for class participation. Students in the class will no doubt have a variety of positions on the ethical issues that we discuss. It is important that you can state a well-informed, well-reasoned argument for your position. Respect for different opinions is important. You will be graded on your classroom participation as well as on the various assignments you are given. This is a class that expects initiative and self-motivation. The instructor cannot possibly tell you everything you need to do to understand or prepare for a certain topic.

You will be exposed to interesting "hot" projects which will have separate project leaders – researchers – coming out of the Dartmouth Experimental Visualization Laboratory (DEVLAB). You need to go out and explore: read books, journals, newspapers, and try to form an informed opinion on the paper or project you are working on. Then bring this information to class to enrich the collective understanding.

You are also expected to turn in your assignment on time and to inform me in advance whether you will have difficulty meeting a deadline, e.g., due to a recruiting trip. Following the **honor principle** is key. This also means that you need to understand how to quote a reference, how to make sure that you give full credit to materials you find elsewhere (e.g., web) and learn how to give due credit to fellow students. Any violation of the honor principle in your paper writing assignments is easily checked and may lead to serious repercussions.

Lateness policy: Points will be deducted if you are late in submitting an assignment. Lateness will result in loss of points that is commensurate to the amount of delay in submission and the excuse behind it. No assignment will be accepted after 5pm the second day it is due without permission.

Reading Materials

Textbook (available at Wheelock Books):

Ethics for the Information Age 1 Book Paperback (limp)
http://www.pearson.ch/pageid/34/artikel/19434AW/Addison-Wesley/0321194349/EthicsfortheInformationAge.aspx
Author: **Michael Quinn**
Publisher: Addison-Wesley

General other Sources:

Book: Ethics & Technology: Ethical Issues in the Age of Information and Communication Technology by Herman Tavani

Book: Ethics in Computing by Kevin Bowyer
Computer Ethics and Social Values, by Deborah Johnson.

Online Materials will be made available at http://www.cs.dartmouth.edu/~cs99 (now still under construction)

Homework Assignments

- Attend class, read the assigned readings and prepare to participate in class discussion
- Write one research papers (3 pages without the references)
- Write second research paper (4 pages without the references)
- Write a third research paper or Report of Final Team Project (6-10 pages depending on the size of the team, for the main report without Appendix and documentation or references)
    - The project will include programming, documentation and a report that explains it. The team is 2-4 persons
- One in-class power point presentation on a topic of your choice, or on one of your papers, or on the final project. You will need to submit an outline-proposal of your presentation first.

Research Papers

First Paper: 3 pages (not counting reference pages).
One carefully presented and analyzed current issue that presents an ethical dilemma or discussion related to computing or computer science. The **ethical discussion is the primary focus in paper one**, while the computer trend or technology is secondary and should be explained in the context of the ethical issue it creates.

Example topics: Stem cell research, medical privacy, a privacy issue such as new legislation on tracking each passenger entering and leaving the country; human cloning and the role of computer science; the use of robotics in war planning and implementation; professors being funded to do military work; energy consumption by industrialized countries; role of computing in upgrading job opportunities for minorities, women or the disabled; free speech and the internet; copying music over the internet; computer science research funded by the department of defense; and many other topics that we can help you choose. Make a clear statement with your

open sentences and narrative cohesive. Keep the topic as focused as possible. We will go over these during first week of classes.

Second Paper: 4 pages (not counting reference pages)

The second papers are on a current trend or cutting edge in computer science of your choice. It can be either an independent paper or preliminary research that prepares the ground for the third paper (team project).

Third paper (or report of team project): 6-10 pages (not counting reference pages and documentation). Your role in the project will be decided as a result of team meetings with the other members and with the project leader who is a Ph.D. graduate student assistant.

Note: Your research papers will be posted on the class website, so they should be written in a technical, non-personal and well-documented format. Research references can range from a minimum of 8 to about 30. Cited references will include the citations within the main text as discussed in class, following scientific convention.

Tentative paper schedule – adjustments may be made to these dates

January 10	**PAPER 1 DRAFT DUE – 5 PM, via email**
January 17	**PAPER 1 FINAL DUE – 5 pm, via email**
January 24	**PAPER 2 DRAFT DUE – 5 PM, via email**
February 7	**PAPER 2 FINAL DUE – 5 PM, via email**
February 18	**PAPER 3 DRAFT DUE (Project proposal emailed to Makedon and corresponding project leader)**
February 28	**SEMIFINAL PAPER 3 VERSION DUE**
March 9	**FINAL PAPER 3 DUE (AND PROJECT)**

Grading

Grades will be assigned as follows (out of 100% total):

20% Paper 1 (ethical dilemma)
25% Paper 2 (technical topic)
35% Paper 3 and Project
10% Presentation
10% Class Participation

Honor Principle

The Dartmouth Honor principle applies to the work in this course. It is expected that the material you hand in is a product of your own work. You are allowed to discuss the cases and issues with others in the class, and with us. If you receive substantial assistance from someone, you should so note on the material you turn in. All sources should be properly acknowledged, according to standard convention (see the Dartmouth booklet Sources:

Their Use and Acknowledgement). For example, you are allowed to quote materials from the technical reports in www.ecomrisk.org, as long as you give proper credit to those references. Materials from websites are to be meticulously annotated by date, page, source, and amount extracted. Website references count as one half of other references.

Disabilities

Students with disabilities, including "invisible" disabilities like chronic diseases, learning disabilities, and psychiatric disabilities should discuss with professor what might be helpful to them. Also, stop by the Academic Skills Center to register for support services.

Projects - Preliminary descriptions

1. Detect and Prevent Malicious Peers in a Peer-to-Peer System

Project Leader: Song Ye, DEVLAB researcher
Website: http://www.cs.dartmouth.edu/~yesong/P2PProject.html
Project Description:
Peer-to-Peer (P2P) is used to identify a class of systems and applications that employ distributed resources to perform some function in a decentralized manner, where every participating node can act as both a client and a server. P2P systems provide some obvious advantages: improved scalability by avoiding dependency on centralized servers, which are often points of failure of the system; elimination of the need for costly infrastructure by enabling direct communication among peers; and easily facilitation of resource aggregation, for example by harnessing available CPU cycles from a large number of peers to provide massive processing power.

A typical P2P system allows mutually distrustful parties with conflicting interests to join or leave freely. Its open, anonymous and decentralized nature makes it extremely difficult to verify the validity of the resources offered by other peers. Furthermore, *free riders*, which only want to use other peers' resources without contributing anything, have greatly compromised the fairness of most P2P systems and discouraged contributing peers from continuing to share resources.

In this project, you are asked to **design a mechanism to detect and prevent malicious peers in a P2P system.** This project is flexible as the students can use any programming language they like and they can choose any P2P systems, existing or imaginary ones. They are asked to define "malicious peers", which is application-specific. More details are available at

---

2. Microarray Gene Analysis: Data Mining in Bioinformatics
Website: http://www.cs.dartmouth.edu/~wyh/microarrayproject.htm
Project Leader: Yuhang Wang, DEVLAB researcher

Project Description:

The project can be done in teams. Each team will be about 4 students and some students do the programming, others write the report, do research, etc.

What is it about?
This project is about data mining in Bioinformatics, and in particular microarray expression data analysis. No previous knowledge of biology beyond highschool level is required. You will find all of the background information in the Resources section below.

This project will involve some programming. You will need to write a program to parse the gene expression data downloaded from public repositories, and to convert it to the ARFF format. You can use Perl or Java, or any other language of your choice. But I would highly recommend Perl because of its superb regular expression capabilities.

Then you are supposed to experiment with the data using WEKA, which is an open source implementation of many state-of-the-art data mining and machine learning algorithms. You can apply feature selection and classification algorithms, and the report the results.

What are the deliverables?
1. Perl scripts and/or Java applications
2. an oral presentation, and
3. a written report.

What to do?
1. Read the materials listed in Resources.
2. Find an interesting gene expression data set in Gene Expression Omnibus (GEO). The data set should have a relatively large number of samples from two or more tissue classes. For example, GDS534 would be an interesting one, which has 75 samples and 3 classes. The orginal paper can be found here.
3. Write the code!
4. Run experiments using WEKA. Observe: Which classifier leads to the best 10-fold cross-validation classification accuracy? If you apply feature selection on the training data prior to classification, do you obtain better accuracy? If so, what is the smallest number of genes that gives you the best accuracy?
5. Write the report.

Resources
1. Wikipedia links for DNA, gene, gene expression, and DNA microarray.
2. A cool Flash Animation explaining DNA Microarray Methodology
3. Home of Weka
4. A brief introduction and overview of Perl is here.
5. Perl regular expressions tutorial.
6. Mark's very short tutorial about references in Perl.
7. A page explaining the ARFF format is here.
8. A page explaining the SOFT data format used in GEO is here.
9.

References

1. Golub, T. R., D. K. Slonim, et al. (1999). "Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring." Science 286(5439): 531-537. [full text]
2. Yuhang Wang, Fillia Makedon, James Ford, and Justin Pearlman. HykGene: A Hybrid Approach for Selecting Marker Genes for Phenotype Classification using Microarray Gene Expression Data. Bioinformatics, in press. [full text]
3. E. Frank, M. Hall, L. Trigg, G. Holmes, and I. H. Witten, "Data mining in bioinformatics using Weka," Bioinformatics, vol. 20, pp. 2479-2481, 2004. [full text]
4. T. Li, C. Zhang, and M. Ogihara, "A comparative study of feature selection and multiclass classification methods for tissue classification based on gene expression," Bioinformatics, vol. 20, pp. 2429-2437, 2004. [full text]

*** For questions, please contact this project's leader: Yuhang Wang(wyh@cs.dartmouth.edu).

---

3. Simulation of an efficient routing algorithm in wireless sensor network
website: www.cs.dartmouth.edu/~cs99/projects/WSNProject.html
Project Leader: Yurong Xu, DEVLAB researcher

Summary of Project:

This project is to do a simulation on routing algorithm for Wireless Sensor Networks. This project will let you understand what is sensor networks, how it works. The project is for team project, which is to be done in teams of about 2-4 students. Students are asked to do some research on it, design and test an algorithm, simulate it, and more.

Motivation Goals

As an emerging new technology, Wireless Sensor Networks(WSN)[1,2,3,4] represent a new type of Ad-hoc networks that integrate sensing, processing, and wireless communication in a distributed system, which is applying in the lots fields of the society, such as surveillance, healthcare, industry and military applications.

Distributed Hash Table technology (DHT)[7][8] is implementing in current P2P networks, and is proved to be a successful technology. But it world be inappropriate to adopt the normal DHT in WSN, while DHT typically interconnect nodes in a way determined by their logical identifications in the DHT, which is largely independent of their proximity in the physical network topology. Such interconnection, in a big possibility, will be impossible in the WSN, because of the limited range of wireless connection.

This project is focused on simulation an efficient routing algorithm[5][6], which is based DHT, but adopted it to WSN with the integration of LST-n (LaST – n connection first) algorithm to get over the disadvantage in the general DHT.

The Deliverables

1. Program
2. Presentation

3. Research Report/Paper

Suggested Methodology

1. Understand what is Sensor Networks, Wireless Sensor Networks[1][2][3].
2. Understand the DHT algorithm in [7][8].
3. Modify the DHT algorithm into WSN
4. Using java or other languages(C or C++) to implement the algorithm
5. Do some small evaluations.

Reference:

[1] I. Akyildiz et al., "A Survey on Sensor Networks", IEEE Commun. Mag., vol. 40, no. 8, Aug. 2002, pp. 102–14.
[2] Ahmed, A.A.; Shi, H.; Shang, Y., "A survey on network protocols for wireless sensor networks", Proceedings of International Conference on Information Technology on Research and Education, ITRE2003, Aug. 2003, pp. 301 – 305.
[3]Vieira, M.A.M.; Coelho, C.N., Jr.; da Silva, D.C., Jr.; da Mata, J.M., "Survey on wireless sensor network devices", Proceedings of IEEE conference on Emerging Technologies and Factory Automation, 2003 (ETFA '03), Sept. 2003, pp.:537 - 544 vol.1.
[4]Al-Karaki, J.N.; Kamal, A.E., "Routing techniques in wireless sensor networks: a survey",IEEE Wireless Communications, Dec. 2004, Pages:6 – 28,Vol.11.
[5]Vincent D. Park , M. Scott Corson, A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks, Proceedings of the INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution, p.1405, April 09-11, 1997.
[6]Sylvia Ratnasamy, Brad Karp, Scott Shenker, Deborah Estrin, Ramesh Govindan, Li Yin, and Fang Yu. Data-centric storage in sensornets with GHT, a geographic hash table. Mob. Netw. Appl., 8(4):427--442, 2003.
[7]Ion Stoica , Robert Morris , David Karger , M. Frans Kaashoek , Hari Balakrishnan, Chord: A scalable peer-to-peer lookup service for internet applications, Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications, p.149-160, August 2001, San Diego, California, United States
[8]G. Manku, M. Bawa, and P. Raghavan. Symphony: distributed hashing in a small world. In 4th USENIX Symp. on Internet Technologies and Systems, pages 127--140, 2003.

---

4. Challenging Cryptography

Website: http://www.cs.dartmouth.edu/~zyle/05cs99Crypto/05cs99Crypto.html

PROJECT LEADER

Zhengyi Le, Researcher, DEVLAB

SUMMARY

Cryptography and network security grow in importance as essential building blocks. The team will choose one state-of-the-art cryptographic algorithm, implement it, and simulate some attacks against it. You may not break it, but you could harm or weaken it on some conditions (or extreme conditions). The experimental results will be analyzed in the written reports.

GOALS

We will use this project to prove the following statement:

With a few exceptions, no crypto algorithms or protocols are proven secure unconditionally.

SUGGESTED METHODOLOGY (DATA SOURCES)
You can use the crypto library (C language) of OpenSSL:
http://www.openssl.org/

Or Java class:
http://java.sun.com/j2se/1.4.2/docs/api/java/math/BigInteger.html

REFERENCES (to get started with)
[1] Bruce Schneier. "Applied Cryptography: Protocols, Algorithms, and Source Code in C." *Second Edition*, *John Wiley & Sons, Inc, 1996*
[2] Dan Boneh and Ramarathnam Venkatesan. "Breaking RSA may not be equivalent to factoring." *Eurocrypt '98, Springer-Verlag LNCS 1233. 1998.*
[3] Dan Boneh**, "**Twenty years of attacks on the RSA cryptosystem", *in Notices of the American Mathematical Society (AMS), Vol. 46, No. 2, pp. 203--213, 1999.*
[4] The Prime Pages: http://www.utm.edu/research/primes/
[5] Electronic Frontier Foundation. *Cracking DES: http://cryptome.org/cracking-des.htm*

---

5. Collaborative Filtering
Website: http://www.cs.dartmouth.edu/~clap/cs99/project.htm
PROJECT LEADER
Sheng Zhang, Researcher, DEVLAB

---

6. TecFlow
Website:TBA
PROJECT LEADER
(By Yan Zhao Yan.Zhao@dartmouth.edu), Researcher, DEVLAB

---

7. Negotiation System
Website:TBA
PROJECT LEADER
(By Yan Zhao Yan.Zhao@dartmouth.edu), Researcher, DEVLAB

Tentative Schedule of Topics Covered
Fillia Makedon
Winter, 2005

Date	*Ethical/Legal Issues*	*Assignments*
Jan. 4 - 6	Introduction to ethics; Course Organization	Ch. 1 and 2 of Quinn (optional: *Tavani 27-61)*

Date	*Ethical/Legal Issues*	*Assignments*
Jan. 11 - 13	Privacy and Encryption	Ch. 5 of text
Jan. 18 - 20	Security	Ch. 6 of Text, (optional: Tavani 152-172)
Jan. 25 - 27	Cybercrime and Cyberterrorism	Ch. 3, 5.10
	Discussion of Security Project	readings for project TBA (optional: Tavani 175-194)
	Regulating Cyberspace; Using Cyberspace for Government Services Tavani 227-255	
Feb. 1 - 3	Intellectual Property	*Ch. 4, (optional Tavani 197-219)*
	Open Source Software	4.8, 4.9, 4.10 (optional: Tavani 220-223)
Feb 8 - 10	Safety Critical Systems	Ch. 7
	and Computer Reliability	(optional Tavani 276-277)
Feb 15 -18	Project preparation meetings with project leaders: Details TBA	
	1.Gene Analysis (Yuhang)	
	2.Crypto Project (Zhengyi)	
	3.Sensor Networks (Yurong)	
	4. Peer to Peer Networks (Song)	
	5. Negotiation of trust (Sheng)	
Feb 22-24	Continue project meetings from above	*Readings TBA for each project*
	Class Presentations	

Date	*Ethical/Legal Issues*	*Assignments*
Mar 1-3	The New Workplace : Fairness; Equity in the Workplace	Ch 8, *(optional Tavani 258-275)*
	Digital Divide, globalization, access	
	Professional Code of Ethics	Ch 9, (optional: *Tavani 87-112)*
	Whistle Blowing	*Ch. 9.6*
Mar. 8-9	Wrap Up and Review	
Mar 10	Final Project Report due	