SHUBHENDU AGGARWAL

Senior Undergraduate
Dept. of Computer Science and Engineering
Indian Institute of Technology, Kanpur, India
E-mail: shubhu@iitk.ac.in , shubhenduaggarwal@gmail.com

G-207, Hall 1
IIT Kanpur-208016
Phone: + (91) 9005751932
http://home.iitk.ac.in/~shubhu

Education

Degree / Certificate	Institution	Year	CPI / Percentage	Rank
Mtech	Indian Institute of Technology, Kanpur, India	2013*		
B.Tech	Indian Institute of Technology, Kanpur, India	2013*	9.0/ 10.0** (3.6 / 4.0)	
Class XII (CBSE)	Lancer's Convent School, New Delhi, India	2008	92.4%	
Class X (CBSE)	Lancer's Convent School, New Delhi, India	2006	93.0%	

*expected year of graduation **after completion of 7th semester

Scholastic Achievements

- Scholarships:
  - Selected for scholarship to attend the workshop organized by the National University of Singapore (NUS) School of Computing (SoC) in May'12 among undergraduate and graduate students from all over Asia
  - Selected for AIEEE scholarship awarded to 300 students from all over India for good performance in AIEEE 2008 examination
- Ranked 341 (99.89 percentile) amongst students from all over India in IIT-JEE 2008
- Ranked 229 (99.97 percentile) amongst students from all over India in AIEEE 2008 and awarded a scholarship
- Ranked 486 (99.69 percentile) amongst students from all over India in GATE 2012
- Awarded the Academic Excellence Awards twice in three years for the years 2010-2011 and 2008-2009 for good academic performance in IIT Kanpur
- Received Merit Certificate in Physics in Class XII CBSE board exams for being amongst top 0.1% students all over India
- Received Certificate of Merit in National Standard Examination in Chemistry (NSEC) 2007-08 and qualified to appear for the Indian National Chemistry Olympiad (Theory) 2008
- Awarded a Certificate of Merit for my performance in Hindi in Class X boards by Hindi Sahitya Academy, Delhi Government
- Awarded Certificate of Merit in the GREEN Olympiad 2004 organised by TERI and Ministry of Environment and Forests, Government of India
- Ranked 46 amongst over 1,50,000 students all over India in the Delhi College of Engineering Common Entrance Examination 2008
- Received Merit Certificate in Physics in Class XII CBSE board exams for being amongst top 0.1 percent students all over India
- Won Third Prize in JK White Design Challenge in Techkriti'10 (annual Technical and Entrepreneurship festival of IIT Kanpur)

Internships / Work Experience

Developing a prototype system to provide Analytics on the Cloud (IBM-IRL, May 2011 – July 2011)

Mentors: Deepak S Padmanabhan, IBM India Research Labs, Bangalore

- Goal is to provide a cloud based service for defining and executing predictive analytics workflows focusing on ease of use, minimal startup cost and flexible pricing based on usage
- A client expecting analytic solutions can upload data sources and choose among configurable pre-defined workflows.
- Configurable workflows: Handles variability in data models by providing entry points corresponding to different stages of data preparation like data merging data and data aggregation.
- The configured workflow runs on IBM SPSS Modeler on the server present on the cloud and gives results back to the customer.
- System was built using JS, Java Servlets, IBM SPSS scripts and Dojo toolkit

The project involved developing a system using which the customer can choose and customize a pre-defined workflow of his choice and get the results of the analysis on the same webpage after the workflow executes in IBM SPSS Modeler on the server present on the cloud.

- Designed pre-packaged workflows for data analysis in the retail sector.
- Built a prototype for providing on-demand analytic solutions for Small and Medium sized Businesses (SMBs).

Problem Addressed: Pre-requisites for usage of Analytics Solutions are purchasing a suite of analytics solutions and high-performance servers to run them on. Many SMBs are cost-conscious and hence unable to afford this initial cost. Solution: We propose providing analytics solutions over the cloud and bill based on usage

Advantages of the solution:

- Scalable on-demand architecture for analytics
- Pay-as-you-go model
- Pre-packaged and configurable workflows
- No high initial cost of setting up server and purchasing a suite of analytical solutions
- No requirement of a analytics expert for operating SPSS suite

Mechanism:

1. Customer uploads data sources on the cloud (if not already present).
2. Customer chooses a workflow depending on his/her requirements.
3. Customer picks up a suitable data entry point in the workflow.
4. Customer chooses a data source (already present on the cloud) for the entry point.
5. Customer repeats this process until all entry points are provided with data sources as per requirement.
6. After providing data sources to all the data entry points of the chosen workflow, customer can provide parameters (if any) required for execution of the workflow through the UI itself.
7. If the attributes of the provided data do not match with those required in the workflow chosen, then the customer will be asked to create a view and match his/her data's attributes to the standard data attributes provided in the workflow. This involves data type checking in the background.
8. After making all the choices, the customer presses the Submit button. The request is transferred to the server on the cloud. SPSS is fired up and the customized workflow is executed.
9. The results are returned to the customer by the server in form of a download link.

Workflow: A workflow is a flow chart representation of the various steps in analytics performed by SPSS

The workflows we have provided in the prototype of our solution are:

1. Unsupervised Customer Segmentation (Clustering)
2. Customer Segmentation – Supervised Learning
3. Association Rule Mining

Based on a study, it was found that the best target for such a system would be the retail sector. These workflows were chosen keeping in mind the requirements of the retail sector.

Implementation:

Java-script, Java servlets, IBM SPSS scripts and Dojo toolkit were used for developing the system. IBM SPSS runs in the background on the server to execute the workflow received from the client and export the results. IBM DB2 is used to store all the data uploaded by the client.

Building a forum for the listeners of a Community Radio Station through a Voice Application (May-July 2010)

Guide: Dr Aaditeshwar Seth, Computer Science Department, IIT Delhi

Goal of this project was to provide a solution to the poor and marginalized communities to create their own local voice media.

- Developed an IVR (Interactive Voice Response) System using Asterisk which is an open source software implementation of a telephone PBX (Private Branch Exchange)
- Recorded messages and caller details in a MySQL database
- An interface was provided for the administrator using JSP (Java Server Pages) and Python

The project is being implemented at Gurgaon Ki Awaaz Samudayik Radio Station in Gurgaon and is actually a part of GramVaani, which is an initiative involved in building innovative technologies to empower poor to voice their opinions and demands.

Developing an interface for MapServer application & understanding working of MapLab with PostgreSQL (June-July 2010)

Internship Guide: Dr P. Chaudhary, Scientist, DTRL, DRDO(Defence R&D Organisation)

MapServer is an open source development environment for building spatially-enabled internet applications and MapLab is a tool

for managing mapping applications and mapfiles.

- Developed an interface to view simple maps from Mapfiles with features like Zooming & Panning using PHP scripting language
- Integrated a PostgreSQL database with MapServer and displayed maps using sample spatial data.

Mtech Thesis

Skyline Queries on High Dimensional Data TO BE UPDATED

Mentor: Prof. Arnab Bhattacharya, CSE IITK and Prof. Sumit Ganguly, CSE IITK

- The goal is to implement skyline queries in high dimensional data for which a large number of skyline points exist.

Conference/Workshop Attended

- Attended talk at **ISEC 2012**, 5th Indian Software Engineering Conference on ***Methods for Loop Invariant Generation:*** *The Fixpoint Brush in The Art of Invariant Generation* by Dr. Sumit Gulwani, Senior Researcher, Microsoft Research (Redmond Lab)

Paper Presentations

- Presented the paper **Cluster Ensemble Selection** by Fern and Lin as part of Data Mining course
- Presented the paper **On *k*-Nearest Neighbor Searching in Non-Ordered Discrete Data Spaces** as part of Indexing and Searching Techniques course.
- Wrote a critique on the paper **Optimizing Matlab through Just-In-Time Specialization** by Boisvert, Hendren, and Verbrugge

Relevant Courses

- Data Mining
- Modern Cryptography
- Operating System
- Compiler Design
- Principles of Database Systems
- Computer Networks
- Principles of Programming Lang.
- Machine Learning
- Algorithms II
- Discrete mathematics
- Theory of computation
- Programming tools & techniques
- Data structures & algorithms
- Intro. to Computer Organization
- Indexing and Searching Techniques
- Advanced Compilers Optimizations
- Multi-agent Systems
- Artificial Intelligence
- Fundamental of computing

Mathematics courses: Mathematical Logic, Multivariable Calculus, Complex Analysis & Linear Algebra, Fourier Analysis & Differential Equations

Economics courses: Microeconomics, Macroeconomics, Indian Problems and Policies, Money and Banking

Technical Skills

Programming Languages	C, C++, Java, Python, Prolog, Oz, Assembly Language
Web Developments	PHP, HTML, JavaScript, MySQL, CSS
Tools	LaTeX, Lex, Yacc, Make, Shell, Eclipse, Matlab, GNU Octave, Autocad
Platforms	Windows, Linux

Key Academic Projects

Indexing and Searching TO BE UPDATED

Multiagent Systems TO BE UPDATED

Advanced Compilers Optimizations TO BE UPDATED

Stock Price Direction Prediction (September-November 2011)

Guide: Dr Krithika Venkataramani, IIT Kanpur

- New indicator attributes were created from the available data to achieve better results.
- Predicted daily stock price direction for Indian stocks using SVMs and decision tree classifiers with an accuracy of around 65%.
- Also predicted weekly trends in prices with an accuracy of around 70%.

- Input data involved opening, closing, high and low daily prices of stock along with volume traded.

Viral Marketing using Social Network Mining (September-November 2011)
Guide: Dr Arnab Bhattacharya, IIT Kanpur
- Viral Marketing is spreading the word about a product from one person to another like a virus. Much more profitable approach to companies than direct marketing.
- A set of most influential people are selected from a social network based on their network value.
- Network value was defined in a novel way so as to incorporate various aspects such as number of friends, strength of the friendships and how influential the friends are.
- Collected social network data from Facebook & Epinions social networks and performed experiments to calculate network value of customers.

Improving time of algorithm for achieving k-anonymity in high dimensional data(September 2011-in progress)
Guide: Dr Arnab Bhattacharya, IIT Kanpur
- Analysed various methods which can be used to achieve improvement in time for achieving k-anonymity in high dimensions while preserving the privacy of data and data integrity.
- Currently, experimenting on a method based on splitting the attribute set and performing some k'-anonymity on each set to achieve an overall k-anonymity.
- Significant improvement in time was observed while following the above approach using Mondrian's algorithm in high dimensions.

Developing a compiler for C (January-April 2011)
Guide: Amey Karkare, IIT Kanpur
- In a team of five, developed a compiler for C to MIPS assembly code using C as implementation language.
- Played major role during coding and in the initial design of the project.
- Important features of C implemented: Multi-dimensional arrays, printf() function, type checking, scoping and function calls apart from the basic C features.

Coding a bot for detecting movable objects in a 2D grid (March-April 2011)
Guide: Dr Harish Karnick, IIT Kanpur
- The bot was coded in Python language to sense its environment and navigate the grid.
- The objects could move with a probability whenever sensed by the bot.
- The bot provided the location of the found objects and the total number of objects in the arena.

Developing a file sharing and P2P video streaming application (August-November 2010)
Guide: Dr Dheeraj Sanghi, IIT Kanpur
- For file sharing, a central server was maintained which maintained file upload and download of various users.
- For video streaming, a tracker was designed which kept data about all available video links among the users connected to the system. A user can demand a video and it will be played on VLC media player at the client as it is streamed from another user.

Cryptanalysis of Popular Cryptosystems (Aug 2011– Nov 2011) (coding in C, Mathematica)
Mentor: Prof. Manindra Agrawal, CSE IITK
- Performed Cryptanalysis of ciphers including permutation & substitution cipher, Variant of DES, AES and RSA
- Cracked 4 round DES using differential cryptanalysis, reduce round AES by Square Attack and 1024 bit RSA with low exponent using LLL attack

Automated Hospital Management System (January-April 2011)
Guide: Dr Sumit Ganguly, IIT Kanpur
- In a group of two, designed and implemented a hospital management system providing easy access to all kinds of data relevant internally to a hospital.
- Provided appropriate views authorised access of data for patients, doctors and receptionist.
- The software managed details like medical stock, locating empty wards, generating bills, medical history of in-patients and doctor's availability.

Simulation and analysis of various features in OS (August-November 2010)
Guide: Dr Sumit Ganguly, IIT Kanpur
- Analysed system calls and scheduling algorithms using OS stimulator NACHOS.
- Implemented and analyzed performance of different mutual exclusion algorithms.
- Simulated and evaluated page replacement algorithms for memory management and a fully associative TLB.

Positions of Responsibility

- **Assistant Coordinator, IOPC** (International Online Programming Contest) in Techkriti'10
    - Assisted in managing more than 1000 teams taking part in the event from around the globe
    - Involved in publicising the event through forums and website
- **English Editor, Vibhuti** (School magazine)
    - Involved in reading, editing and selecting a large number articles from students to be published in the magazine

Extra Curricular Activities/Interests

- Participated in several inter school quizzes, amongst the best being reaching the finals of the prestigious Annual Science Quiz organized by the National Science Centre, Delhi in 2005.
- Won Third Prize in JK White Design Challenge in Techkriti'10 (annual Technical and Entrepreneurship festival of IIT Kanpur)
    - Found out an eco-friendly and innovative usage of white cement using bamboo fibres and made several samples after collecting data
- Participated in the workshop and was awarded Certificate of Merit in the GREEN Olympiad 2004 organised by TERI and Ministry of Environment and Forests, Government of India
- Applied my technical know-how to design, implement and put in the field an innovative system which helped illiterate villagers to communicate with each other and create their own voice based local media through a voice application system of mobile phones
- Interested in playing volleyball and cricket
- Part of the cricket team in playing several inter-club matches in Delhi during the years 2004-2006
- Have a published newspaper article on "Does violence in cinema have an adverse affect on children" in the Times of India in the year 2004