ANURAG MALIK

5852008743 · am3926@g.rit.edu
github.com/anuragmalik

EDUCATION

Master of Science, Computer Science Expected May-2018
Rochester Institute of Technology, NY. GPA: 3.4/4.0

TECHNICAL SKILLS

Programming	Java, C++, Python, SQL (MySql)
Data Science	Machine Learning, Big data analysis and visualization - scikit-learn, Pandas
Distributed Systems	Distributed Key-Value Store, Distributed Cache, Apache Spark, Map Reduce
Parallel Programming	GPU & CUDA Programming

EXPERIENCE

Graduate Teaching Assistant & Grader *Rochester Institute of Technology, NY* Jan 2016 - May 2017

· Assisting in designing course content, conduct recitations, mentoring & teaching the undergraduate course on basic algorithms and data structures in Python & Java. Grader for 'Computational Problem Solving' graduate level course.

Software Development Engineer *Sopra Steria Pvt. Ltd., India* Jun 2013 - Jun 2015

· Assisted in design, Java development & maintenance activities for multi-asset finance product of the Sopra Banking Suite. Completed 21 client requests and product development tasks under agile project management.

· Discovered several inefficiencies in legacy code and worked on SQL optimization, performance analysis, and profiling, and developed external data structure library for efficient implementation.

PERSONAL PROJECTS

Orbox-C++ : Single Player Puzzle Game April 2017

· A single player puzzle game with manual and auto-solver mode, developed using an object oriented design in C++14 including usage of clone design pattern, STL and Boost Library. Also implemented a recursive backtracking solver for three different levels in the game.

GPU implementation of All Source Shortest Path Algorithms using CUDA Dec 2016

· Implemented parallel versions of various 'All source shortest path' algorithms on large graphs. Implemented using *CUDA on GPU* & *Parallel Java 2 library* on CPU. Compared weak & strong scaling results.

Highly Available Distributed Key-Value Store Jul 2016

· Amazon DynamoDB like scalable & fault tolerant key-value store with *consistent hashing on virtual nodes, fault tolerant hinted-handoff replication* & *Merkel trees* for anti-entropy.

ACADEMIC PROJECTS & RESEARCH

Rental Listings: Data analysis on apartment listings dataset May 2017

· This project aims at solving a classification problem in apartment rental listings data corpus and develop a system capable of evaluating all listings based on their quality. Implemented supervised learning techniques and build decision tree model. System is evaluated using the multi-class logarithmic loss.

Parallel K-Means clustering using MapReduce Dec 2016

· Implemented and evaluated a parallel k-means clustering algorithm using *MapReduce parallel programming* technique. The algorithm can scale well and efficiently process large datasets on commodity hardware.

Network reconstruction using distributed UAVs *Research* Aug 2016

· Worked on implementing distributed area exploration algorithms for UAVs based on sparse connectivity probing & finding best locations for relay deployment in reconstructing network holes.

Chord - Distributed File System May 2016

· A scalable distributed file system built using *JavaRMI*, with consistent hashing over multiple nodes capable of handling file search, download and upload requests from clients.

Restaurant Recommendation System & RESTful APIs Apr 2016

· Responsive web application and sentiment analysis engine build using *Naive Bayes classifier with chi-square filtering* on crowd-sourced reviews developed with *Bootstrap* & *Jersey, JAX-RS* on *MongoDB* database.