



## SPECIALITIES

AWS Certified | Data Analytics | Machine Learning | Big Data | IoT | Computer Networks | Distributed Systems

## TECHNICAL SKILLS

<b>Languages:</b>	Java, Python, C#/.Net, C/C++, R, JavaScript, Octave, MATLAB
<b>Libraries:</b>	Java-DOM, log4j, Python - Pandas, NumPy, SciPy, Scikitlearn, PyMongo
<b>OS:</b>	Linux, Windows, Bash/Batch scripting
<b>Software IDE</b>	Visual Studio, NetBeans, Eclipse, Lidlipse, Spyder
<b>Software Platforms:</b>	AWS, Hadoop, Pig, Hive, MS SQL Server & MongoDB, Tableau
<b>Software Tools:</b>	Git, Jenkins, Docker, Agile methodology

## WORK EXPERIENCE

**Software Developer**      **Red Hen Systems, Fort Collins, CO**      **Nov 2016 – Current**

- Responsible for development of IsWhere™, Media Geo-Tagger™ software, based on C#/.Net.
- Introduced online licensing & management system for all our software products, which gives a seamless approach to manage licenses and has reduced the time between licensing and accessing the product.
- Improved the software by adding tools to clip videos, modified UI, and provided compatibility for different file types - Experience with windows forms and WPF.
- Batch automation process to copy video files, decode and extract data, creation of structured data files such as CSVs/XMLs, ArcPy processing (.shp).
- Interacted with clients, involving GIS projects, to gather requirements and manage workflows.
- Optimization of Methane gas classifier – Several hundred thousand spatially referenced data involved.
- Created a website to access client data online. ArcGIS server & MS SQL server is on the backend.
- Followed agile methodology for code review, collaboration, presentation and final inclusion to product.

**Grad. Research Asst.**      **CNRL, Colorado State University, Fort Collins, CO**      **Feb 2015 – May 2017**

Reconstruction of the Topology of Undirected Graphs from Partial Information - Python

- Used extended Robust PCA, an accurate method for matrix completion by low rank approximation.
- Applied on 2D & 3D WSN's and real world internet & social networks with 500 - 5000 nodes.
- Facebook, E-mail, Collaboration, CAIDA etc. are some of the real world online/social networks.
- Reconstructed the topology of WSNs embedded in 2D & 3D physical spaces with mean error less than 20% for 80% missing information in virtual topological coordinate space.
- Reconstructed the social networks with less than 10% information giving a mean error less than 6% & 0.5 hops.

**Software Programmer**      **Fromme Custom Solutions, Fort Collins, CO**      **May 2015 – Jul 2015**

- Developed a software application to test storage networks for SMI specs.
- The storage networks are tested for the compliance of TLS protocol versions & cipher suites.
- Application created in Java uses DOM & log4j libraries. Outputs .xml and .csv files.

## EDUCATION

<b>Master of Science</b>	Computer Engineering	<b>GPA: 3.50/4.00</b>	<b>May 2017</b>
Colorado State University			

---

## PROJECTS

---

### **Building a Distributed, Replicated, and Fault Tolerant File System - Java**

- Implemented a Distributed File System in Java. The system comprises of Controller, Chunk Server & Client.
- The files are stored in chunks of 64KB with a replication level of 3. The system adds fault tolerance by constantly checking and fixing errors.

### **Scalable Server design using Thread pools and load balancing on the server - Java**

- Designed and implemented Threadpools to handle load on the server, without using 3rdparty libraries.
- Observed that mean per client throughput X No of active connections equals server throughput. Std. Deviation of per client throughput is low.

### **Twitter Sentiment Analysis – MapReduce – Python (OAuth)**

- Calculate sentiment score for each tweet. Derive inferences & sentiment for new words using score of the tweet.

### **Train Support Vector Machine model to detect and classify attacks – Python & Sci-kit learn**

- Used SVM SVC - Varied C & gamma to obtain accuracy and F1 score. An accuracy of 98% was achieved.
- Cross-validated the model for binary class with 5 folds. The dataset contains one million sample & 40 features.

### **Blog Creation with NoSQL database MongoDB – Python (PyMongo and Bottle)**

- Created a functional site, with MongoDB database in backend, for storing articles, comment and login details in of a blog. Text based indexes for faster search.

### **Estimating PageRank Values of Wikipedia Articles using MapReduce – Java, Apache Hadoop**

- Implemented page rank algorithm on AWS, to rank the internal Wikipedia articles with Wiki data dump.

### **Content Searching in a Distributed Application layer Network – Structured & Unstructured - Java**

- Implemented Unstructured & Structured (Chord Algorithm) P2P network on an 80-node test bed.
- Network formation and file searching on UDP and TCP.

### **K Means Clustering and Dimensionality Reduction(Principal Component Analysis) - Octave**

- Image compression implemented with K-means. The image is compressed with 16 color clusters.
- Dimensionality reduction is achieved by choosing K principal components in PCA. Achieved 10x reduction.

### **Recognition of Handwritten digits(0-9) using Logistic Regression & Neural networks - Octave**

- Implemented a basic one-vs-all classifier with multiple regularized logistic regression classifiers.
- Also implemented a neural network scheme. Computed the performance by measuring accuracy.

---

## PUBLICATIONS

---

- Capture and Reconstruction of the Topology of Undirected Graphs from Partial Coordinates: A Matrix Completion Based Approach, **Thesis/Dissertation – Colorado State University**.
- Network Topology Mapping from Graph Geodesics and Partial Virtual Coordinates, submitted **IEEE/ACM Trans**.
- Topology Maps and Distance Free Localization from Partial Virtual Coordinates for IoT, **IEEE ICC 2016**.