Jingwen Shi

Personal Website | Linkedin | shijingwen9@gmail.com | (+1) 517-974-8921 | East Lansing, MI 48823

EDUCATION

Focus: Network Security, AI for Distributed Systems, Cloud Computing

- Ph.D., Computer Science and Engineering, Michigan State University — *Aug. 2019 - May 2024*
- M.S., Computer Technology, University of Chinese Academy of Sciences — *Sept. 2016 - May 2019*
- B.S., Information Security, Hunan University, Graduated with Honor (Summa eq.) — *Sept. 2012 - May 2016*

INTERNSHIP

Research Intern, Los Alamos National Lab — *June 2021 - Aug. 2021*

- Developed a learning approach to re-construct a black-box cyber-physical system model from scratch by observing operating data samples. The cyber-physical system consists of physical equations and control rules.
- Applied (1) **ordinary least squares** regression to infer the variables and corresponding coefficients of physical equations; (2) **support vector machine** to infer the decision hyperplane of control rules. Reduced redundant parameters by **singular value decomposition**. Attained an accuracy of **97%**.

Research Intern, Alibaba — *Jan. 2019 - June 2019*

Project1. Traffic Prediction and Uncertainty Estimation for Resource Allocation in Cloud

- Queries per Second (QPS) prediction is critical for resource allocation in the cloud. Developed **Bayesian Neural Networks** to dynamically predict the upper bounds of QPS, where the upper bounds are more crucial than the mean predictions.
- Fetched data from the distributed monitoring system **EagleEye** and trained the **deep learning** model on **Hadoop**. One paper accepted to [JST'19].

Project2. Anomaly Detection for Virtual Machines Failure in Cloud

- Developed an efficient **anomaly detection** framework reducing the false alarm rate in detecting virtual machine failure due to the inherent jitters in low-level system metrics(e.g., I/O, CPU usage).
- Fetched low-level system metrics logs from large-scale clusters in parallel and persisted the data to the local machines using **Shell** and **SQL**. Integrated Machine learning and statistic models in a hierarchical pipeline from multi-views using **isolation forest**, **three-sigma**, and **kernel density estimation (KDE)**. Reduced **95%** of false alarms without sacrificing the true alarm rate.

PROJECTS

Uncovering Insecure Designs of Cellular 911 Services — *Sept. 2021 - Dec. 2021*

- Built a **4G LTE testbed** consisting of the core network (**OpenIMSCore, MySQL**, and **Linux**) and Radio Access Network (**srsRAN**, **USRP**). Successfully defended **DoS attacks** on 911 services using **TLS** encryption. One paper accepted to [Mobicom'22].

Uncovering Insecure Designs of Cellular Voice Services — *Jan. 2020 - May 2021*

- Discovered design loopholes in the PDCP layer of the **radio protocol stack** and QoS from **3GPP** standards.
- Developed (1) an **overshadow attack** that can mute voices during the call by modifying **srsRAN**; (2) an **identity linkage attack** that can discover users' radio network identities; (3) a **free-charging attack** that can bypass the charging policy rules of operators.

A Big Data System for Geographical Data — *Feb. 2016 - April 2016*

- Implemented a data pipeline to store raw data in **HDFS** and connect HDFS with **PostgreSQL** of a cloud system for storing, querying, and analyzing large-scale satellite images and traffic data.
- Designed a spatial index approach that can reduce I/O traffic by up to **70%**.
- Built **Hbase** and **MongoDB** for performance evaluation. One paper accepted to [IPCCC'18].

PUBLICATIONS

- **Network Security**: [TMC'22, Paper], [Mobicom'22, Paper], [Mobisys'23 Peer Reviewing]
- **AI and Data Privacy**: [AAAI'23 Peer Reviewing, Paper]
- **Distributed System**: [IPCCC'18, Paper], [JST'19, Paper]

SKILLS

- **Frameworks**: srsRAN, OpenIMSCore, Android; MongoDB, Hbase, PostgreSQL, Spark, Hadoop; Django, D3.js, Scrapy
- **Languages**: Python, Java, C, Julia, Matlab, PHP, JavaScript, Shell, SQL
- **Developer Tools**: Linux, Git, Github