Akond Rahman

Current Address:
2941 Ligon Street Apartment D23, Raleigh, North Carolina, USA
+1 919 758 1347
aarahman@ncsu.edu
http://akondrahman.github.io/

Research Interests

Software engineering focused in the area of DevOps, Infrastructure as Code, and Mining Software Repositories

Education

- **Doctor of Philosophy (Ph.D.) in Computer Science** Aug 2014 - Present
  - Adviser: Dr. Laurie Williams
  - North Carolina State University, Raleigh, NC, USA
- **Master of Science (M.Sc.) in Computer Science and Engineering** Jan 2012 - May 2014
  - University of Connecticut, Storrs, CT, USA
- **Bachelor of Science (B.Sc.) in Computer Science and Engineering** Dec 2004 - Sep 2009
  - Adviser: Dr. Mahmuda Naznin
  - Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

Awards & Honors

- **Microsoft Open Source Challenge**
  I am the **Grand Prize** winner of the 2016 Microsoft Open Source Challenge. I used Microsoft Research's Deep Semantic Similarity Model (DSSM) tool to quantify the semantic similarity of software repositories. Details are available here http://tiny.cc/unp1by.
- **First prize in IUT National ICT Festival**
  In August, 2009, I received this award in IUT National ICT Festival, in Software Projects category.
- **Dean's List Award**
  I received this award for achieving academic excellence for two semesters in the year of 2008 during my Bachelors program in Bangladesh University of Engineering and Technology.
- **University Scholarship for Merit**
  I received this award for maintaining decent academic progress for four semesters from 2007 to 2009 during my Bachelors program in Bangladesh University of Engineering and Technology.

Professional Experience

- **Data Science Intern** May 2017 - Aug 2017
  - IBM Watson and Cloud, Research Triangle Park (RTP), NC, USA
  - Activities:
    - Mining open source and industry software repositories to assess the merits of continuous integration on open source and proprietary software projects.
    - Applied statistical tests and modeling techniques on software meta data.
    - Collaborated with full time members and interns of the IBM Watson and Cloud team.
- **Systems Infrastructure Intern** May 2016 - Aug 2016
  - Redhat Inc., Raleigh, NC, USA
  - Activities:

    - Developed monitoring facilities for the OpenShift infrastructure to facilitate continuous deployment of internal Redhat applications.
    - Facilitated Integration of in-house Redhat application into the OpenShift infrastructure.
    - Collaborated with full time members and interns of the PlatOps and SysEng team at Redhat.

- **Research Assistant** Jan 2016 - May 2016
  - Science of Security Lablet, National Security Agency (NSA), USA
  - Activities:
    - Investigated research aspects of DevOps.
    - Research methods for science of security.
    - Collaboration with Science of Security Lablet participants.

- **Software Research Intern** May 2015 - Aug 2015
  - ABB Corporate Research, Raleigh, NC, USA
  - Activities:
    - Designed, and developed a software framework to detect similar software applications in ABB.
    - Collaborated with full time members and interns of the Software Architecture Team and Industrial Software Engineering team at ABB Corporate Research.

- **Teaching Assistant** Aug 2014 - Dec 2015
  - Department of Computer Science, North Carolina State University, Raleigh, NC, USA
  - Activities:
    - Provided lectures for a graduate course.
    - Manged Apache Virtual Computing (VCL) images to administer course projects and homeworks.
    - Designed requirements for class projects.
    - Designed questions for class quizzes, and exams for graduate and undergraduate courses.
    - Graded homeworks, and exams for graduate and undergraduate courses.

- **Graduate Assistant** Jun 2013 - May 2014.
  - Engineering Computing Services (ECS), University of Connecticut (UConn), Storrs, CT, USA
  - Activities:
    - Troubleshooted ECS computer systems and services.
    - Solved day to day software and hardware problems of the administrative staff in School of Engineering at UConn.

- **Software Engineer** Jan 2010 - Jun 2011
  - Dohatec New Media, Dhaka, Bangladesh
  - Activities:
    - Designed and developed a Java-based multi-modal biometric system using the MegaMatcher SDK.
    - Developed and tested an online procurement system for the Ministry of Planning of the People's Republic of Bangladesh using ASP.NET with C# and Microsoft SQL Server.

Analytical Skills

Natural Language Processing: Deep Structured Semantic Model (DSSM), N-gram Models, Topic Model (LDA)

Optimization: Differential Evolution (DE), MaxWalkSat, Simulated Annealing

Qualitative Analysis: Grounded theory, Survey analysis and design

Regression: Linear and Logistic Regrssion

Sampling: Oversampling, Synthetic Minority Over-sampling Technique(SMOTE), Undersampling

Statistics: ANOVA, Association Rule Mining, Chi-square Test, Correlation Analysis, Effect Size, Inter-rater Reliability, Non-parametric Hypothesis Tests, Principal Component Analysis

Supervised Learning: Artificial Neural Network, Decision Tree, Hidden Markov Model, kNN Classifier, Maximum Likelihood Classifier, Naive Bayes Classifier, Random Forest, Support Vector Machine

Unsupervised Learning: DBScan, Hierarchical Clustering, K-Means Clustering

Languages and Tools

Programming Languages: Bash, C, C#, CSS, HTML, Java, PHP, Python, R, SQL

Data Analysis Tools: Apache Spark with Python, Natural Language Toolkit (NLTK), R Studio, Scikit-learn

Database Tools: Microsoft SQL Server (v. 2005, 2008), MySQL (v. 5.0, 5.5), PostgreSQL

DevOps Tools: Amazon Web Services (EC2, Code Deploy), Ansible, Git, Jenkins, Puppet, Vagrant

Development Kits: Android SDK, Plugin Development for Chef and Puppet

Peer Reviewed Publications

- **Akond Rahman**, and Laurie Williams, "Characterizing Defective Configuration Scripts Used for Continuous Deployment", in 11th International Conference on Software Testing, Validation, and Verification (ICST '18), 2018, *to appear*.
- **Akond Rahman**, Priysha Pradhan, Asif Partho, and Laurie Williams, "Predicting Android Application Security and Privacy Risk with Static Code Metrics", in 4th International Conference on Mobile Software Engineering and Systems (MOBILESoft '17), 2017, pages: 149-153.
- **Akond Rahman**, Asif Partho, David Meder, and Laurie Williams, "Which Factors Influence Practitioners' Usage of Build Automation Tools?", in 3rd International Workshop on Rapid Continuous Software Engineering (RCoSE '17), 2017, pages: 20-26.
- Morgan Burcham, Mahran Al-Zyoud, Jeffrey Carver, Mohammed Alsaleh, Hongying Du, Fida Gilani, Jun Jiang, **Akond** Rahman, Ozgur Kafali, Ehab Al-Shaer, and Laurie Williams, "Characterizing Scientific Reporting in Security Literature: An analysis of ACM CCS and IEEE S&P Paper", in Symposium and Bootcamp on the Science of Security (HotSoS'17) 2017, pages: 13-23.
- **Akond Ashfaque Ur Rahman**, "Code Metrics For Predicting Risk Levels of Android Applications", in Proceedings of 2016 KSU Conference on Cybersecurity Education, Research and Practice (CCERP' 2016), October, 2016.
- **Akond Ashfaque Ur Rahman** and Laurie Williams, "Software Security in DevOps: Synthesizing Practitioners' Perceptions and Practices", in Proceedings of International Workshop on Continuous Software Evolution and Delivery (CSED), May, 2016, Austin, TX, USA.
- **Akond Ashfaque Ur Rahman** and Laurie Williams, "Security Practices Used in DevOps", in Proceedings of Symposium and Bootcamp on the Science of Security (HotSoS), April, 2016, Pittsburg, PA, USA.
- **Akond Ashfaque Ur Rahman**, Eric Helms, Laurie Williams, and Chris Parnin, "Synthesizing Continuous Deployment Practices in Software Development", in Proceedings of 13^{th} Agile Conference, pages 1-10, Washington D.C., USA, August, 2015.
- **Akond Ashfaque Ur Rahman**, Md. Atiqul Islam Mollah, and Mahmuda Naznin, "Multiple Targets Tracking Using Kinematics in Wireless Sensor Networks" in Wireless Sensor Network, pages 263-274, August, 2011.
- **Akond Ashfaque Ur Rahman**, Mahmuda Naznin, and Md. Atiqul Islam Mollah, "Energy Efficient Multiple Targets Tracking Using Target Kinematics in Wireless Sensor Networks" in Proceedings of 4^{th} International Conference on Sensor Technologies and Applications (SensorComm), pages 275-280, Venice, Italy, July, 2010.
- **Akond Ashfaque Ur Rahman**, Mahmuda Naznin, and Md. Atiqul Islam Mollah, "Service Priority Based Target Tracking Framework in a Wireless Sensor Network" in Proceedings of 3^{rd} IEEE International Conference on Computer Science and Information Technology (ICCSIT), pages 389-392, Chengdu, China, July, 2010.
- M.M.Shahiduzzaman, Mahmuda Naznin, and **Akond Ashfaque Ur Rahman**, "Portable and Secure Multimedia Data Transfer in Mobile Phones Using Record Management Store (RMS)" in Proceedings of 3^{rd} IEEE International Conference on Computer Science and Information Technology (ICCSIT), pages 364-367, Chengdu, China, July, 2010.

Presentations

- "Predicting Android Application Security and Privacy Risk with Static Code Metrics", in 4th International Conference on Mobile Software Engineering and Systems (MOBILESoft '17), Buenos Aires, Argentina, May, 2017 URL

- "Which Factors Influence Practitioners' Usage of Build Automation Tools?", in 3rd International Workshop on Rapid Continuous Software Engineering (RCoSE '17), Buenos Aires, Argentina, May, 2017 URL

- "Software Security in DevOps: Synthesizing Practitioners' Perceptions and Practice", in First International Workshop on Continuous Software Evolution and Delivery, Austin,TX, USA, May, 2016 URL

- "Synthesizing Continuous Deployment Practices in Software Development", in 13^{th} Agile Conference, Washington D.C., USA, August, 2015 URL

Major Research Projects

- **On The Merits of Continuous Integration for Proprieatary Projects**:

  Continuous integration (CI) tools integrate code changes by automatically compiling, building, and executing test cases upon submission of code changes. Use of CI tools is getting increasingly popular, yet how proprietary projects reap the benefits of CI remains unknown. To investigate the influence of CI on software development, we mine 661 open source software (OSS) projects, and 171 proprietary projects. For OSS projects, we observe the expected benefits after CI adoption, i.e. more bugs are resolved, and more issues are resolved. However, for the proprietary projects, we cannot make similar observations. Therefore, we cannot claim that CI is the 'silver bullet' for software development. Our findings indicate that only adoption of CI might not be enough to improve software development. CI can be effective for software development if practitioners use CIâĂŹs feedback mechanism efficiently, by applying the practice of making frequent commits. For proprietary projects we observe practitioners to commit less frequently, and hence not use CI effectively, for obtaining feedback on the submitted code changes. We recommend practitioners to (i) apply the CI best practices along with adoption of CI tools, (ii) consider their teamâĂŹs development context before adopting CI tools, and (iii) after adoption of CI, investigate if CI satisfies their needs by applying software analytics.

- **Anti-patterns in Infrastructure as Code Scripts**:

  Defects in infrastructure as code (IaC) scripts can have serious consequences for organizations who adopt DevOps. By identifying which characteristics of IaC scripts correlate with defects, we can identify anti-patterns, and help software practitioners make informed decisions on better development and maintenance of IaC scripts, and increase quality of IaC scripts. We focus on characteristics of IaC scripts and IaC development that correlate with defects. For our initial studies, we mined open source version control systems from three organizations: Mozilla, Openstack, and Wikimedia, to identify the defect-related characteristics and conduct our case studies. From our empirical analysis, we identify (i) 14 IaC code and four churn characteristics that correlate with defects; and (ii) 12 process characteristics such as, frequency of changes, and ownership of IaC scripts that correlate with defects. Using the identified characterestics we build prediction models to predict which script is likely to be defective.

- **Mining Discussion Topics related to Infrastructure as Code (IaC) from Q&A Websites**:

  Infrastructure as Code (IaC) tools are integral to DevOps. As a relatively new topic, IaC programmers can face issues with technologies, and identifying these issues can help in idnetifying the knowledge base, as well as improvize the tools themselves. In this research project I investigate the topics that IaC programmers are facing from five question and answer websites. I use topic modeling and qualitative analysis to identify these topics. For topic modeling I use LDA, and perplexity to identify the optimal number of topics used in case of topic modeling.

- **What Factors Influence the Usage of Build Automation Tools?**:

  Modern release engineering is a big shift from traditional software engineering with respect to development and deployment practices. In this research project I investigate the factors that influence the usage of modern release engiineering software tools such as Ansible, Git, Maven, and Puppet. As part of conducting this project I am collecting survey responses, and analyzing those using regression techniques. The ultimate goal of this project is to systematically find and quantify the factors that affect the usage of

build automation tools. I am using logistic regression and association rule mining to identify the factors that signficantly influence usage of build automation tools.

- **Predicting Android Application Security and Privacy Risk Using Code Metrics**:

  Prior research has provided empirical evidence on how Android applications can expose security and privacy issues of Android users. In this research project I aim to investigate if code metrics such as number of functions, cyclomatic complexity can be used to predict multiple levels of risk for Android applications. In this project I have used multiple supervised and unsupervised machine learning techniques such as decision tree, gaussian naive bayes classifier, hierarchical clustering, k-means clustering, knn classifier, random forest, SMOTE, and support vector machine.

- **Investigating the Influence of Context Switching on Programmer Productivity**:

  In this project I have analyzed a dataset generated by Codealike, a plugin that tracks programmer activities from popular IDEs such as Visual Studio, and Eclipse, and reports developer activity accordingly. In this research project I am investigating to categorize 84,030 developer programming sessions that are spread across 15,178 days. Another goal of this research project is to investigate the effects of context and task switching on programmer productivity measured in terms of programmer actions such as code, debug, and build.

- **Perception and Adoption of Software Security in DevOps Organizations**:

  The goal of this ongoing research project is to identify the software engineering practices used for software security in IT organizations who have adopted DevOps. We have studied 350 Internet artifacts to gain background knowledge about security aspects of DevOps. Using this background knowledge we would like to identify the notion of software security among different IT organizations who have adopted DevOps, and what software practices they are using.

- **Supporting Enterprise-Wide Software Reuse Using Semantic Similarity of Software Projects**:

  In this ongoing project, I along with my colleagues at ABB Corporate research are investigating how semantic similarity amongst software projects can be quantified, and used to perform software reusability. To detect software similarity, our tool called *Simila*, converts software applications into natural language tokens. These natural language tokens are used as input to *semantic comparator*, to compare the natural language tokens of one project to another. As next steps, first we are planning to detect software similarity amongst open source project repositories, and then amongst enterprise software repsoitories.

- **Investigating Continuous Deployment Practices Used in Software Development**:

  The goal of this research project was to aid software practitioners in implementing continuous deployment through a systematic analysis of software practices that are used by software companies. I studied the continuous deployment practices of 19 software companies by performing a qualitative analysis of Internet artifacts and by conducting follow-up inquiries. In total, I found 11 software practices that are used by 19 software companies. I also found that in terms of use, eight of the 11 software practices are common across 14 software companies URL: `http://www.realsearchgroup.org/realsearch/agile-software-development/`.

Major Professional Projects

- **Integration of In-house Applications and Monitoring in OpenShift**:

  This project focuses on integrating monitoring alerts so that OpenShift hosts can be monitored for troubleshooting and debugging. The project also included integration of *Redhat Insights* into the Openshift architecture. I implemented the project using Ansible, Python and Puppet scripts.

- **SIMILA: Detecting Similar Software Applicatons in ABB**: I developed a software framework that detects similarity amongst software applications by converting software artifacts into natural language tokens. I implemented the project using SrcML.NET, Swum.NET, and Python Natural Language Toolkit (NLTK). To facilitate performance gain and scalability I used Apache Spark with Python.

- **Multi Modal Biometric SDK**: I developed a software framework that provides multi-functional biometric utilities such as capturing, quality checking, auto rotation, cropping, etc. of biometric images like face, finger, and iris. I implemented this project using the Megamatcher 3.1 SDK tool.

- **Central Procurement Management Information System (PROMIS)**: I was actively in development, and testing of this software package which is a information management system developed for the Planning Ministry of the Government of People's Republic of Bangladesh. For implementation I used C#, and Microsoft SQLServer 2005.

Voluntary Experience

- **Student volunteer**: I was a student volunteer for the following conferences:
  - 39^{th} International Conference on Software Engineering (ICSE) 2017
- **SubReviewer**: I was a sub reviewer for the following conferences:
  - 28^{th} International Symposium on Software Reliability Engineering (ISSRE) 2017
  - 11^{th} International Symposium on Empirical Software Engineering and Measurement (ESEM) 2017
  - 14^{th} International Conference on Mining Software Repositories (MSR) 2017
  - 37^{th}, and 40^{th} International Conference on Software Engineering (ICSE) 2015, 2018
- **Team Leader for NCSU Team at IT Architecture Competition**:
  I led the NCSU team in the Student Showdown competition arranged by the Raleigh Chapter of International Association for Software Architects (IASA) (`http://members.iasaglobal.org/news/281134/Iasa-IT-Architect-Competition.htm`).

Major Course Projects

- **Optimizing Integrated Project Model Defect Flow Chains for Software Development Teams**:

  The goal of this project is to help software engineering practitioner to analyze the Integrated Project Model Defect Flow Chain proposed by Abdel-Hamid and Madnick, by implementing the model, and optimizing the model using a standard optimizer. In this project I implemented a standard software model that takes development effort into account and calculates the amount of development errors with certain assumptions, and considering a subset of the full development resource inputs. To optimize the development resources needed to reduce development errors we use differential algorithm (DE). URL: `https://github.com/akondrahman/59115ASE/tree/master/project`.

- **Automatic Deployment of Applications on Amazon EC2 Using DevOps Practices**:

  The goal of this project was to automatically deploy Java-based web applications using two DevOps practices namely, *blue green deployment*, and *chaos monkey*. In this project I implemented the concept of blue-green deployment that re-routes traffic based on monitoring a certain application. I also implemented a miniature version of *chaos monkey* a pratice that Netflix uses for reliability testing. I used Amazon EC2 service, Amazon Code Deploy, Java Melody, and Vagrant to implement this project. URL: `https://github.com/akondrahman/Miscellaneous/tree/master/DevOps`.

- ***Decaf* Compiler**: The goal of this project was to implement a compiler for *Decaf*, a subset of the Java programming language. Me and my partner used Java, JavaCC, and Jasmin to compile and run 120 Decaf programs, to implement this project.