

AKOND RAHMAN

2941 Ligon Street Apartment D23, Raleigh, North Carolina, USA
aarahman@ncsu.edu
<https://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
 - Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
-

AWARDS & HONORS

- **ACM SIGSOFT Doctoral Symposium Award**
My thesis proposal related to DevOps and Infrastructure as Code won the ACM SIGSOFT Best Doctoral Symposium Award at ICSE 2018, Gothenburg, Sweden.
 - **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>.
 - **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

- **Research Intern** May 2018 - Aug 2018
 - IBM Research, IBM TJ Watson Research Center, Yorktown Heights, NY, USA
 - Activities:
 - * Integrating AI Marketplace protocol with Hyperledger Composer and Fabric.
 - * Application development with blockchain technologies such as Hyperledger Composer and Fabric.
- **Data Science Intern** May 2017 - Aug 2017
 - IBM, 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 project metadata.

- **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.
 - **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.
 - **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.
 - **Research Assistant** Jan 2016 - May 2016
 - Science of Security Lablet, National Security Agency (NSA), USA
 - Activities:
 - * Investigated attack surface modeling to find vulnerabilities.
 - * Studied research methods for science of security.
-

ANALYTICAL SKILLS

Natural Language Processing: Bag-of-Words (BOW), Term Frequency Inverse Document Frequency (TF-IDF), Deep Structured Semantic Model (DSSM), Topic Model (LDA)

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

Qualitative Analysis: Grounded theory, Survey analysis and design

Regression: Linear and Logistic Regression

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

Statistics: ANOVA, Association Rule Mining, Chi-square Test, Correlation Analysis, Effect Size, Interrater Reliability, Non-parametric and Parametric Hypothesis Tests, Principal Component Analysis (PCA)

Supervised Learning: Artificial Neural Network, Decision Tree, 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: Apache Spark with Python, Natural Language Toolkit (NLTK), R Studio, Scikit-learn

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

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

Security: Findbugs, Fortify, IBM BlueMix Static Analyzer, Microsoft Threat Modeling Tool (MSTMT)

Blockchain: Hyperledger Composer, Hyperledger Fabric

PEER REVIEWED PUBLICATIONS

- **Akond Rahman**, Amritanshu Agrawal, Rahul Krishna, and Alexander Sobran, “Characterizing The Influence of Continuous Integration: Empirical Results from 250+ Open Source and Proprietary Projects”, in 4th International Workshop on Software Analytics (SWAN 2018), co-located with European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE) 2018, *to appear. Pre-print*
- **Akond Rahman**, Jonathan Stallings, and Laurie Williams “Poster: Defect Prediction Metrics for Infrastructure as Code Scripts in DevOps”, in Companion Proceedings of the 40th International Conference on Software Engineering (ICSE '18), 2018, *Pre-print*
- **Akond Rahman**, Asif Partho, Pat Morrison, and Laurie Williams “What Questions Do Programmers Ask About Configuration As Code?”, in Proceedings of the 4th International Workshop on Rapid Continuous Software Engineering (RCoSE '18), co-located with International Conference for Software Engineering (ICSE) 2018, Gothenburg, Sweden, May, 2018, *Pre-print*
- **Akond Rahman**, “Comprehension Effort and Programming Activities: Related? Or Not Related?”, in International Conference of Mining Software Repositories (MSR) 2018, Gothenburg, Sweden, May 2018., *Pre-print*
- **Akond Rahman**, “Anti-patterns in Infrastructure as Code”, in International Conference of Software Testing, Validation, and Verification (ICST-PhD Symposium) 2018, Vasteras, Sweden, April 2018, *Pre-print*
- **Akond Rahman**, “Characteristics of Defective Infrastructure as Code Scripts in DevOps”, in Companion Proceedings of the 40th International Conference on Software Engineering (ICSE '18), 2018, *Pre-print*
- Rahul Krishna, Amritanshu Agrawal, **Akond Rahman**, Alexander Sobran and Tim Menzies, “What is the Connection Between Issues, Bugs, and Enhancements? (Lessons Learned from 800+ Software Projects)”, in Companion Proceedings of the 40th International Conference on Software Engineering (ICSE '18), 2018, *Pre-print*
- Amritanshu Agrawal, **Akond Rahman**, Rahul Krishna, Alexander Sobran and Tim Menzies, “We Don't Need Another Hero? The Impact of ”Heroes” on Software Development”, in Companion Proceedings of the 40th International Conference on Software Engineering (ICSE '18), 2018, *Pre-print*
- **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, *Pre-print*
- **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 Ashfaq 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 Ashfaq 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 Ashfaq 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 Ashfaq Ur Rahman**, Eric Helms, Laurie Williams, and Chris Parnin, “Synthesizing Continuous Deployment Practices in Software Development”, in Proceedings of 13th Agile Conference, pages 1-10, Washington D.C., USA, August, 2015.
- **Akond Ashfaq 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 Ashfaq Ur Rahman**, Mahmuda Naznin, and Md. Atiqul Islam Mollah, “Energy Efficient Multiple Targets Tracking Using Target Kinematics in Wireless Sensor Networks” in Proceedings of 4th International Conference on Sensor Technologies and Applications (SensorComm), pages 275-280, Venice, Italy, July, 2010.
- **Akond Ashfaq Ur Rahman**, Mahmuda Naznin, and Md. Atiqul Islam Mollah, “Service Priority Based Target Tracking Framework in a Wireless Sensor Network” in Proceedings of 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), pages 389-392, Chengdu, China, July, 2010.
- M.M.Shahiduzzaman, Mahmuda Naznin, and **Akond Ashfaq Ur Rahman**, “Portable and Secure Multimedia Data Transfer in Mobile Phones Using Record Management Store (RMS)” in Proceedings of 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), pages 364-367, Chengdu, China, July, 2010.

PRESENTATIONS

- “What Questions Do Programmers Ask About Configuration As Code?”, in 4th International Workshop on Rapid Continuous Software Engineering (RCoSE’18) URL
- “Comprehension Effort and Programming Activities: Related? Or Not Related?”, in International Conference of Mining Software Repositories 2018 (MSR’18) URL
- **Akond Rahman**, “Characteristics of Defective Infrastructure as Code Scripts in DevOps”, in International Conference on Software Engineering 2018 (ICSE’18) URL
- “Characterizing Defective Configuration Scripts Used for Continuous Deployment”, in 11th International Conference on Software Testing, Validation, and Verification 2018(ICST’18) URL
- “Predicting Android Application Security and Privacy Risk with Static Code Metrics”, in 4th International Conference on Mobile Software Engineering and Systems 2017(MOBILESoft’17) URL

- “Which Factors Influence Practitioners’ Usage of Build Automation Tools?”, in 3rd International Workshop on Rapid Continuous Software Engineering 2017(RCoSE ’17) URL
 - “Software Security in DevOps: Synthesizing Practitioners’ Perceptions and Practice”, in First International Workshop on Continuous Software Evolution and Delivery 2016(CSED) URL
 - “Synthesizing Continuous Deployment Practices in Software Development”, in 13th Agile Conference 2015 (Agile2015) URL
-

VOLUNTARY EXPERIENCE

- **Student volunteer:** I was a student volunteer for the following conferences:
 - 39th International Conference on Software Engineering (ICSE) 2017
 - 40th International Conference on Software Engineering (ICSE) 2018
 - **Sub-reviewer:** I was a sub reviewer for the following conferences:
 - 9th International Symposium on Engineering Secure Software and Systems (ESSoS) 2018
 - 4th International Workshop on Rapid Continuous Software Engineering (RCoSE) 2018
 - 1st International Workshop on Security Awareness from Design to Deployment (SEAD) 2018)
 - 28th International Symposium on Software Reliability Engineering (ISSRE) 2017
 - 11th International Symposium on Empirical Software Engineering and Measurement (ESEM) 2017
 - 14th International Conference on Mining Software Repositories (MSR) 2017
 - 37th, and 40th International Conference on Software Engineering (ICSE) 2015 and 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 PROFESSIONAL PROJECTS

- **Towards Trusted AI Marketplaces Using Blockchain:** As part of my internship at IBM Research I implemented a protocol that implements the concept of AI marketplaces, where consumers and data providers can exchange data and models in a trusted fashion using blockchain. I integrated the protocol with two blockchain systems Hyperledger Fabric and Hyperledger Composer, and observed Fabric-based implementation to be faster compared to that of the Composer-based implementation.
 - **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.
-

MAJOR RESEARCH PROJECTS

- **Security Anti-patterns in Infrastructure as Code Scripts:**
As IaC scripts hold crucial information about the deployment environment, violation of security objectives can be disastrous. We refer to characteristics of IaC scripts that violate security objectives as security

anti-patterns. As an example anti-pattern, if administrator credentials are hard-coded in IaC scripts, attackers can use those credentials and hack into the deployment infrastructure. Systematic investigation can help in identifying which security anti-patterns occur in IaC scripts and at which frequency.

- **On The Merits of Continuous Integration for Proprietary 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 CI's 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.

- **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 identifying the knowledge base, as well as improvise 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.

- **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.

- **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.

- **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/>.