

Md Masudur Rahman

CONTACT INFORMATION Department of Computer Science University of Virginia Charlottesville, VA - 22904, USA +1 434 466 7304 masud@virginia.edu <http://mmasud.me/>

RESEARCH INTERESTS My research interest focuses on applying Information Retrieval and Machine Learning techniques on large-scale software repositories to analyze and automate software engineering systems.

- Software Engineering, Code Search, Mining Large Scale Software Repository, Information Retrieval, Deep Learning, Big Code

EDUCATION **University of Virginia**, Charlottesville, VA, USA

Ph.D., Computer Science, *Started:* August, 2015

- CGPA: 3.95 (up to Fall 2017)

Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh

B.Sc., Computer Science and Engineering, February 2013

- Major: *Theoretical Computer Science*
- CGPA: 3.84 (*Ranked 6th out of 142 students in class*)

PUBLICATIONS

- Evaluating How Developers Use General-Purpose Web-Search for Code Retrieval. **Md Masudur Rahman**, Jed Barson, Sydney Paul, Joshua Kayan, Federico Andres Lois, Sebastian Fernandez Quezada, Christopher Parnin, Kathryn T. Stolee, Baishakhi Ray. IEEE/ACM 15th International Conference on Mining Software Repositories (MSR 2018), 11 pages, 2018. (to appear)
- Which Similarity Metric to Use for Software Documents? A study on Information Retrieval based Software Engineering Tasks. **Md Masudur Rahman**, Saikat Chakraborty, Baishakhi Ray. The 40th International Conference on Software Engineering (ICSE '18 Companion), 2 pages, 2018. (to appear)
- Finding Similar Projects in GitHub using Word2Vec and WMD. **Md Masudur Rahman**. Workshop on the Naturalness of Software (NL+SE 2016) at FSE 2016. ([Slides](#))
- Topic Model based Privacy Protection in Personalized Web Search. Wasi Ahmad, **Md Masudur Rahman**, Hongning Wang. The 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR2016), 4 pages, 2016. ([Link](#))

RESEARCH
EXPERIENCE

Graduate Research Assistant
Supervisor: Baishakhi Ray

Summer 2016 - current

Research Project

- **A study on Information Retrieval based Software Engineering Tasks**

Evaluated the performance of different similarity metrics (VSM, BM25, LSI, and embedding based WMD) on various SE documents w.r.t. two tasks: (i) similar project recommendation and (ii) bug localization. It has been observed that in general the context-aware models achieve better performance on textual artifacts. In contrast, simple keyword-based bag-of-words models perform better in code artifacts.

- **Finding Similar Projects on GitHub**

Leveraged Word2Vec word embedding and Word Mover Distance (WMD) to measure the similarity of GitHub projects utilizing projects' source code and available textual description.

- **Privacy in Information Retrieval**

Proposed a client-centered approach to hide user intents that prevent search engine from building user profile by injecting cover queries with the original user-issued query in a controlled way. Each user's query is submitted with a set of cover queries which carry a similar amount of information to generalize user profile while still having some personalization.

TEACHING
EXPERIENCE

Graduate Teaching Assistant

Fall 2015 - Spring 2017

Department of Computer Science, University of Virginia

Teaching Instructor

Spring 2013 to Summer 2015

Department of Computer Science and Engineering, BRAC University

WORK
EXPERIENCE

- Lecturer

February 2013 - August 2015

Department of Computer Science and Engineering

Brac University, Dhaka, Bangladesh

TECHNICAL
SKILLS

Python, Java, C/C++, Apache Lucene, Solr.

ONLINE
PROFILE

LinkedIn: <https://www.linkedin.com/in/masud99r>

GitHub: <https://github.com/masud99r>

REFERENCE Baishakhi Ray
Assistant Professor
University of Virginia
Email: rayb@virginia.edu
Phone: (434) 982-2212
Fax: (434) 982-2214
Website: <http://rayb.info/>