

Shehab Sarar Ahmed

Website: <https://shehab003.github.io/>

Email: [shehab@cse.buet.ac.bd](mailto:shehab@cse.buet.ac.bd)

[\[Google Scholar\]](#) [\[GitHub\]](#)

## PRESENT STATUS

---

- **Bangladesh University of Engineering and Technology (BUET)** Dhaka, Bangladesh  
*Lecturer, Department of CSE, BUET* October 21, 2019 - Present

## EDUCATION

---

- **Bangladesh University of Engineering and Technology (BUET)** Dhaka, Bangladesh  
*M.Sc. in Computer Science and Engineering; CGPA: 4.0/4.0* April 2019 - Present
- **Bangladesh University of Engineering and Technology (BUET)** Dhaka, Bangladesh  
*B.Sc. in Computer Science and Engineering; CGPA: 3.99/4.0* February 2015 - April 2019  
*\*Ranked 1st in a class of 150 students*

## WORK EXPERIENCE

---

- **Bangladesh University of Engineering and Technology (BUET)** Dhaka, Bangladesh  
*Part-time Lecturer, Department of CSE, BUET* May, 2019 - October, 2019
- **Samsung R&D Institute Bangladesh** Dhaka, Bangladesh  
*Junior Software Engineer* April, 2019 - May, 2019

## RESEARCH EXPERIENCE

---

- **Protein Succinylation Site Prediction using Amino Acid Features**

Protein post-translational modifications (PTM) are important cellular regulatory processes that occur after protein synthesis. Protein PTMs have been detected by a variety of experimental techniques. Generally, the experimental analysis of PTMs requires time-consuming, labor- and capital-intensive techniques and the use of hazardous/expensive chemical reagents. Succinylation, a special type of PTM, involves the transfer of a succinyl group to a protein molecule at its lysine (K) residue. Succinylation plays an important role in the growth and development of plants and is very dynamic under different cellular conditions. In our study, we define a context window for each lysine residue containing its neighboring amino acid residues and apply state-of-the-art NLP models to classify the lysine residue as succinylated or non-succinylated. We also apply meta heuristics to find out optimum set of [amino acid indices](#) to achieve better accuracy. We have compared our work with the state-of-the-art work by using the PLMD dataset and found comparable results with significantly less computational cost. We are now working on further improvement of our experimental result.

**Supervisor:** [Dr. M. Sohel Rahman](#), CSE, BUET

**Status:** Work ongoing

- **Hierarchical Attention for Host Intrusion Detection**

The host-based intrusion detection system (HIDS) has been gaining attention in the community of cybersecurity for over two decades. Compared with the network-based intrusion detection system (NIDS), HIDS has the superiorities of fine granularity and the ability to detect internal attacks. In our work, we propose a novel hierarchical attention based deep learning method to detect intrusion on a host. We evaluate our model on ADFA-LD dataset, which is a collection of Linux system calls' traces. We have tuned our model's hyper parameters to gain higher sensitivity as well as lower false alarm rate compared to other relevant state-of-the-art works.

**Supervisor:** [Dr. Shohrab Hossain](#), CSE, BUET

**Status:** Close to completion

- **Characterization of intrinsically disordered regions in proteins informed by human genetic diversity**

All proteomes contain both proteins and polypeptide segments that don't form a defined three-dimensional structure, yet are biologically active—called intrinsically disordered proteins and regions (IDPs and IDRs). Most of these IDPs/IDRs lack useful functional annotation limiting our understanding of their essentiality for organism fitness. Here we characterized IDRs using protein sequence annotations (or features) of functional sites and regions available in the UniProt knowledgebase (active site, ligand-binding pocket, regions mediating protein-protein interactions, etc.). By measuring the statistical burden of twenty-five features in 981 IDRs of 561 human proteins, we identified eight features that are generally enriched in IDRs. We then collected the genetic variant level data from the general population and patient-based cohorts and evaluated the prevalence of pathogenic and population variations in IDPs/IDRs. We observed that some IDRs tolerate 2 to 12-times more single amino acid substituting missense mutations than synonymous changes. Additionally, we quantified that 37% of all germline pathogenic mutations in 96 IDPs are located in disordered regions of these proteins. Comparing the observed to expected frequency of mutations, we categorized 34 IDRs in 20 IDPs (DDX3X, KIT, RB1, etc.) as intolerant to genetic inactivation upon mutations. Finally, using statistical analysis and a machine learning-based approach, we demonstrate that mutation-intolerant IDRs carry a distinct signature of functional features. Our study presents an orthogonal approach to assign functional importance to IDRs, i.e., by leveraging the wealth of genetics data, which will aid in a deeper understating of the role of IDRs in biological processes and disease mechanisms.

**Supervisor:** [Sumaiya Iqbal](#) & [Dr. Sohel Rahman](#) **Status:** Under review at PLOS Computational Biology

## PUBLICATIONS

---

- Sumaiya Iqbal, Eduardo Pérez-Palma, Jakob B. Jespersen, Patrick May, David Hoksza, Henrike O. Heyne, **Shehab S. Ahmed**, Zaara T. Rifat, M. Sohel Rahman, Kasper Lage, Aarno Palotie, Jeffrey R. Cottrell, Florence F. Wagner, Mark J. Daly, Arthur J. Campbell, Dennis Lal, "Comprehensive characterization of amino acid positions in protein structures reveals molecular effect of missense variants", Proceedings of the National Academy of Sciences of the United States of America, 2020. [\[Link\]](#)
- Sumaiya Iqbal, David Hoksza, Eduardo Pérez-Palma, Patrick May, Jakob B. Jespersen, **Shehab Sarar Ahmed**, Zaara Rifat, Henrike O. Heyne, M. Sohel Rahman, Jeffrey R. Cottrell, Florence F. Wagner, Mark J. Daly, Arthur J. Campbell, Dennis Lal, "MISCAST: Missense variant to protein StruCture Analysis web SuiTe", Nucleic Acids Research, 2020. [\[Link\]](#)
- Sumaiya Iqbal, Jakob B. Jespersen, Eduardo Pérez-Palma, Patrick May, David Hoksza, Henrike O. Heyne, **Shehab S. Ahmed**, Zaara T. Rifat, M. Sohel Rahman, Kasper Lage, Aarno Palotie, Jeffrey R. Cottrell, Florence F. Wagner, Mark J. Daly, Arthur J. Campbell, Dennis Lal, "Burden analysis of missense variants in 1,330 disease-associated genes on 3D provides insights into the mutation effects", Cold Spring Harbor Laboratory, 2020. [\[Link\]](#)
- **Shehab Ahmed**, Zaara Rifat, Arthur J. Campbell, A. Keith Dunker, Sohel Rahman, Sumaiya Iqbal, "Burden of Functional Features and Genetic Variations in Human Intrinsically Disordered Proteins", Biophysical Journal, 2020. [\[Link\]](#)
- Zaara Rifat, **Shehab Ahmed**, Arthur J. Campbell, A. Keith Dunker, Sohel Rahman, Sumaiya Iqbal, "Map of Genetically Constrained Regions in Human Intrinsically Disordered Proteins", Biophysical Journal, 2020. [\[Link\]](#)
- Sumaiya Iqbal, Jakob B. Jespersen, Eduardo Pérez-Palma, Patrick May, David Hoksza, Henrike O. Heyne, **Shehab S. Ahmed**, Zaara T. Rifat, M. Sohel Rahman, Kasper Lage, Aarno Palotie, Jeffrey R. Cottrell, Florence F. Wagner, Mark J. Daly, Arthur C. Campbell, Dennis Lal, "Insights into protein structural, physicochemical, and functional consequences of missense variants in 1,330 disease-associated human genes", Cold Spring Harbor Laboratory, 2019. [\[Link\]](#)

## RESEARCH INTEREST

---

- Network Security
- Natural Language Processing
- Bioinformatics

## TEACHING EXPERIENCE (SELECTED)

---

- Computer Graphics
- Microprocessors, Microcontrollers, and Embedded Systems
- Operating System
- Compiler
- Database
- Data Structure and Algorithm
- Object Oriented Programming Language

## TECHNICAL SKILL

---

- **Programming Languages:** Python, R, C, C++, Java, Assembly Language (8086), PL/SQL
- **Database:** mySQL, Oracle, PostgreSQL, Neo4j
- **Frameworks:** Pytorch, Keras, Trax, Tensorflow
- **Others:** JMeter, Firebase, Google Apps Script

## EXTRA CURRICULAR ACTIVITIES

---

- **Member of Organizing Committee**  
International Conference on Networking Systems and Security (NSysS), organized by Department of CSE, BUET (2019, 2020)
- **Subreviewer**  
Protein Inference Algorithmics: the Peptide Quantity Assignment Model, submitted in COCOON 2021
- **Coach**  
BUET International Collegiate Programming Contest (ICPC) Teams (2019, 2020)
- **Member**  
Departmental Tender Evaluation Committee (2020-up to date)

## HONORS AND AWARDS (SELECTED)

---

- Dean's Award in each academic year in BUET for academic result
- University Merit Scholarships in each semester in BUET for academic result
- Medals in National and Regional Math and Informatics Olympiads

## ACADEMIC PROJECTS (SELECTED)

---

- Calculator built with iGraphics
- Desktop Chat app built using JavaFX
- Android app that manages house owners and tenants [[Video Link](#)].
- Sonar ball game using Atmega32 [[Video Link](#)].
- Cancer detection from x ray images of lungs.

## REFERENCE

---

- **Dr. M. Sohel Rahman**  
Professor, Department of CSE  
Bangladesh University of Engineering and Technology (BUET)  
Email: [msrahman@cse.buet.ac.bd](mailto:msrahman@cse.buet.ac.bd)  
[\[Google Scholar Profile\]](#)
- **Dr. Md. Shohrab Hossain**  
Professor, Department of CSE  
Bangladesh University of Engineering and Technology (BUET)  
Email: [mshohrabhossain@cse.buet.ac.bd](mailto:mshohrabhossain@cse.buet.ac.bd)  
[\[Google Scholar Profile\]](#)
- **Dr. A. B. M. Alim Al Islam**  
Professor, Department of CSE  
Bangladesh University of Engineering and Technology (BUET)  
Email: [alim\\_razi@cse.buet.ac.bd](mailto:alim_razi@cse.buet.ac.bd)  
[\[Google Scholar Profile\]](#)