Data Science and Cyber Security Systems: A Statistical Perspective

Associate Professor Azizur Rahman, PhD
Data Science Research Unit (DSRU)
School of Computing and Mathematics
Charles Sturt University
Australia

Email: azrahman@csu.edu.au

Abstract:
Data science is empowering many contemporary innovations, including the development of robust cybersecurity systems. It is becoming a significant research area that embedded the combination of modern computing tools and statistical inferences, with the demand of adequately maintaining data privacy within the systems. For example, in e-healthcare systems, IoT devices monitor the patient's health and upload classified data as Electronic Medical Records (EMRs) to the cloud. Sharing EMRs data in hyper-connected systems with a limited storage capacity of IoT devices in a third party based service provider is more vulnerable to attack by sophisticated cybercriminals. Existing privacy protection schemes, including cryptography, cannot ensure the privacy demands in a cloud environment. Medical data sharing need multiple data accessing paradigms with different levels of privacy according to the need for consultation by a specialist, e.g., surgeons, general practitioners, allied health staff. Besides, none of these techniques is robust in privacy preservation because of not determining privacy weight (i.e. an orthogonal design matrix-based weight generation strategy for determining data privacy) for the e-Health industry using a cloud environment or automated by adaptation of AI. These effectively limit the full potential of the dynamic cloud-based e-healthcare systems and business environment and the wider uptake and comprehensive benefits of data science more generally.

In this inaugural talk, I will address some significant computing and inferential challenges in data science with a particular focus on cyber design space, security of data, computation, and inferences. It will not only highlight computing statistics for the decision-making process or running intelligent algorithms, but also there must be a focus on reliability measures on any applied outputs. This keynote presentation will also illustrate some real-world case studies from data science, including a quick overview of one of our project aims to develop a robust solution for privacy-preserving e-health data sharing in a cloud environment with an AI-based automated technique to ensure data privacy and measuring reliability.

Short bio:

[Azizur Rahman, PhD](#) is an applied statistician and data scientist with expertise in both developing and applying novel methodologies, models and technologies. He is the Leader of "Statistics and Data Mining Research Group" within the Data Science Research Unit at Charles Sturt University, Australia. He is able to assist in understanding multi-disciplinary research issues within various fields including how to understand the individual activities which occur within very complex scientific, behavioural, socio-economic and ecological systems.

Prof. Rahman develops "alternative methods in microsimulation modelling technologies" which are very useful tools to socioeconomic policy analysis and evaluation. He has more than 110 publications including a few books. His [2020](#) and [2016](#) books have contributed significantly to the fields of "data science and policy analysis" and "small area estimation and microsimulation modelling" respectively. His research is funded by the Australian Federal and State Governments, and he serves on a range of editorial boards including the International Journal of Microsimulation (IJM) and Sustaining Regions. He received several awards including the SOCM Research Excellence Award 2018 and the CSU-RED Achievement Award 2019.