A Communication Efficient Distributed Algorithm for Regression Analyses with Applications to Studying Drug Safety Outcomes

Yong Chen, Ph.D.
Associate Professor of Biostatistics
Senior Scholar at Institute of Biomedical Informatics
Department of Biostatistics, Epidemiology and Informatics
University of Pennsylvania

Multicenter study has advantages in studying rare events or conditions, such as adverse events associated with new medications, association of disease with a rare gene variant, and many others. To overcome the barrier of patient-level data sharing, we propose privacy-preserving and communication-efficient distributed algorithms for regression analyses. The performance of the algorithms is investigated through simulation study and application to datasets from two large distributed networks, including a study of risk factors of acute myocardial infarction (acuteMI) using claims data from five different databases. The results demonstrate that the proposed algorithms are communication-efficient while providing relatively accurate estimation when studying rare drug safety outcomes.