

6<sup>th</sup> International Conference on

# BIOSTATISTICS AND BIOINFORMATICS

November 13-14, 2017 | Atlanta, USA

***En-Bing Lin****Central Michigan University, USA*

## Big data analysis in bioinformatics

With the increasing use of advanced technology and the exploding amount of data in bioinformatics, it is imperative to introduce effective and efficient methods to handle Big data using the distributed and parallel computing technologies. Big data analytics can examine large data sets, analyze and correlate genomic and proteomic information. In this presentation, we begin with an overview of Big data and Big data analytics, we then address several challenging and important tasks in bioinformatics such as analyzing coding, noncoding regions and finding similarities for coding and noncoding regions as well as many other issues. We further study mutual information-based gene or feature selection method where features are wavelet-based; the bootstrap techniques employed to obtain an accurate estimate of the mutual information and other new methods to analyze data. Given the multi-scale structure of most biological data, several methods will be presented to achieve improvements in the quality of mathematical or statistical analysis of such data. In a DNA strand, it is essential to find sequences, which can be transcribed to complementary parts of the DNA strand. We will mention several methods to identify protein coding regions. We also use some special variance and entropy to analyze similarities among coding and noncoding regions of several DNA sequences respectively and compare the resulting data. We will address the use of big data analytics in many phases of the bioinformatics analysis pipeline.

## Biography

En-Bing Lin is a Professor of Mathematics at Central Michigan University, USA. He has been associated with several institutions including Massachusetts Institute of Technology, University of Wisconsin-Milwaukee, University of California, Riverside, University of Toledo, UCLA, and University of Illinois at Chicago. He has received his PhD from Johns Hopkins University. His research interests include Data Analysis, Applied and Computational Mathematics, and Mathematical Physics. He has Supervised a number of graduate and undergraduate students. He serves on the Editorial Boards of several journals. He has organized several special sessions at regional IEEE conferences and many other professional meetings.

[lin1e@cmich.edu](mailto:lin1e@cmich.edu)

## Notes: