

Special Issue on **Biological Big Data Problems: Proposed Solutions and Future Directions**

CALL FOR PAPERS

Our growing power to automatically collect data from natural phenomena continues to surprise us in several science areas. Just to give a few examples, consider satellites, telescopes, and genome sequencers. Novel technologies allow gathering in a single month more data nowadays than what was obtained in previous years or even decades of continuous observation. It is not simple to store, recover, and process a plethora of data using conventional database management systems and algorithms.

Big data suffers from very known computational problems concerning computational time and space complexity. For instance, a homology report under a metagenomic perspective is nothing less than intimidating. Also, considering the sequencing process of a metagenomic project, the redundant genomic reads, a desirable behavior to assure base quality, forces a significant space complexity on storing, sharing, and transmitting the data. We still cannot find solutions that are open source and easy to use applied to the biological domain; leading companies only offer them at the cost of robust computing.

A convenient approach to solving particular big data problems is to devise heuristic solutions. Heuristic solutions have the advantage of avoiding brute-force search. Also, there is a lack of consensus in deciding what programming languages should be used to manipulate big data. Some high-level programming languages can easily produce desirable code at the cost of efficiency; others can be implemented more efficiently at the expense of longer development time. Additionally, unsupervised machine learning algorithms play a significant role in big data analytics since many relationships among data need to be elucidated and cannot be assumed from the onset.

We invite researchers to contribute original articles as well as review articles that will illustrate and stimulate the effort to reduce time and space complexity on big data analytics under the biological context. Also, we would like to create a debate about advantages and hurdles of programming languages to manipulate big data, besides encouraging the availability and development of learning algorithms for biological problems.

Potential topics include but are not limited to the following:

- ▶ Data Mining via unsupervised learning
- ▶ Scalable distributed databases for big data management
- ▶ Heuristic methods to reduce time and space complexity
- ▶ Programming languages applicable to manipulation of big data
- ▶ Parallel programming libraries for big data processing

Authors can submit their manuscripts through the Manuscript Tracking System at <http://mts.hindawi.com/submit/journals/scientifica/genomics/bbdp/>.

Lead Guest Editor

Anderson Santos, Federal University of
Uberlândia, Uberlândia, Brazil
santosardr@ufu.br

Guest Editors

Preetam Ghosh, Virginia
Commonwealth University, Richmond,
Virginia, USA
pghosh@vcu.edu

Marcelo Maia, Federal University of
Uberlândia, Uberlândia, Brazil
marcelo.maia@ufu.br

Fernanda A. Dorella, Federal University
of Minas Gerais, Belo Horizonte, Brazil
fernandadorella@gmail.com

Manuscript Due

Friday, 5 May 2017

First Round of Reviews

Friday, 28 July 2017

Publication Date

Friday, 22 September 2017