Parallel Software Package for The Construction and Analysis of Complex Networks

Hisham Ihshaish and Johan Dijkzeul

VORtech B.V. - Delft (The Netherlands)

"LINC"
learning about interacting networks in climate

VORTECH BV
THE SCIENTIFIC SOFTWARE ENGINEERS

Abstract

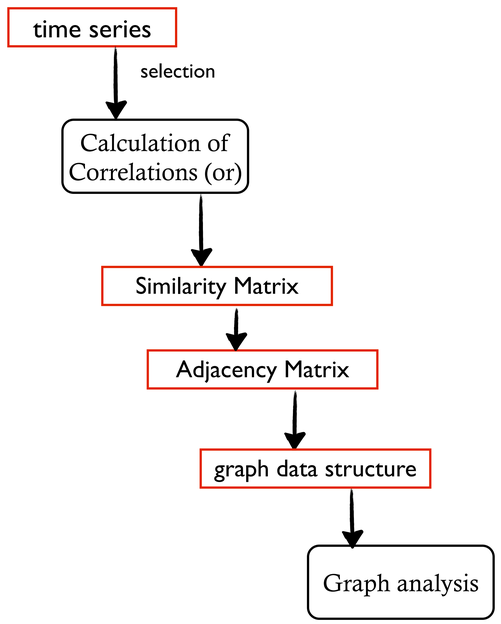
In climate research, big and complex networks could be generated by the big climate data produced by high resolution climate models, and also observations. To analyze such complex networks, there are two main computational challenges concerning both the construction and the analysis of these complex networks:

- the construction of the network, as such, the size of the on-chip physical memory of single processor computing machines might limit the possibility to efficiently construct such big networks.
- computational time needed to analyze complex networks stands to be another real challenge.

High-performance software tools are needed to process both steps (construction and analysis) efficient.

A description of the parallel software package which had been developed in LINC project is presented here.

Processing Climate Networks



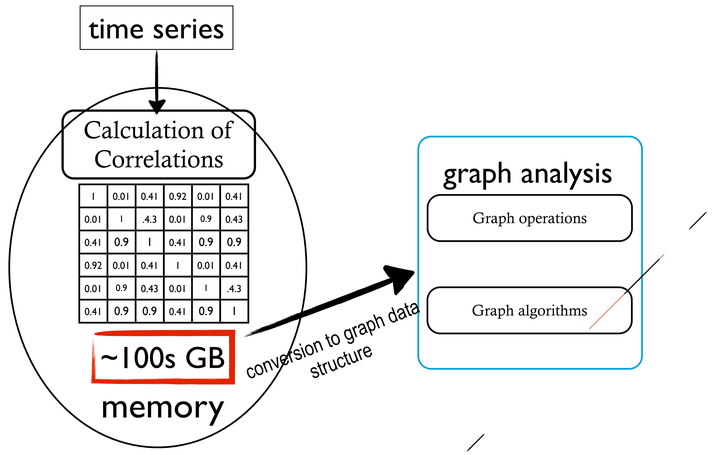
Input / Output: big data (100s of GB).

Processing: time consuming calculations in function of algorithmic complexity and graph type and size.

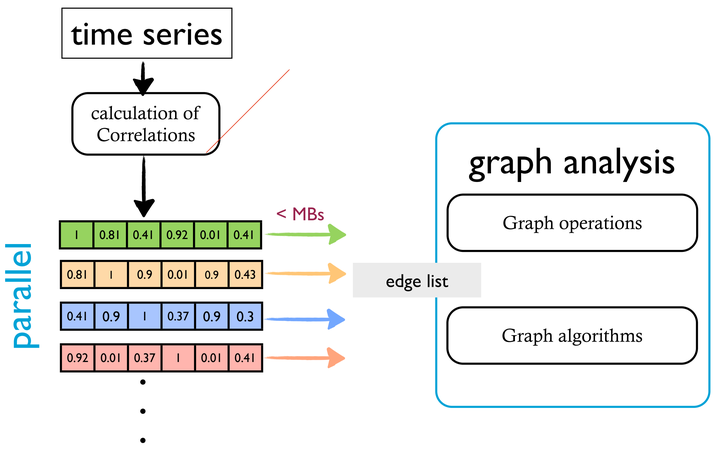
Parallel software tools for network construction

Parallel processing techniques are applied to fasten the process of network construction from large sets of time series (hundreds of thousands to millions). As such, the calculation of the similarity matrix (for instance), corresponding to each node is done independently, so that the correlation of each chunk of nodes is processed in parallel.

A Memory wall for large matrices



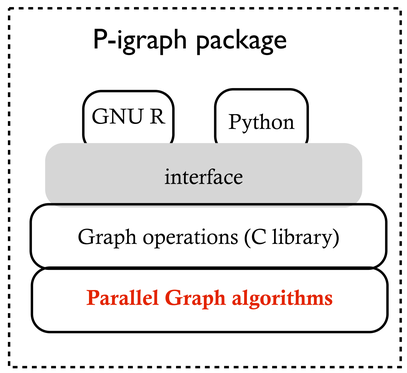
B Parallel calculation of similarity matrices, and transformation to graph data structures of less size (edge list, adjacency list)



Parallel software tools for network analysis

Igraph library was paralleled applying the shared-memory parallel programming model OpenMP. Most of the implemented algorithms in Igraph library are now parallel.

A P-igraph architecture



B Parallel software package for complex networks analysis and construction.