

 HELSINKI UNIVERSITY OF TECHNOLOGY
 LABORATORY OF COMPUTER AND INFORMATION SCIENCE

Research Group: Parsimonious modeling

<http://www.hiit.fi/pm>


Jaakko Hollmén
*Lab. of Computer and Information Science,
 Helsinki University of Technology*


 HELSINKI UNIVERSITY OF TECHNOLOGY
 LABORATORY OF COMPUTER AND INFORMATION SCIENCE

Parsimonious Modeling

- Group leader: Jaakko Hollmén
- Graduate students (5): Salla Ruosaari, Mika Sulkava, Jarkko Tikka, Mikko Korpela, Janne Toivola
- Home page: <http://www.hiit.fi/pm>


Jaakko Hollmén: Parsimonious modeling research group 2


 HELSINKI UNIVERSITY OF TECHNOLOGY
 LABORATORY OF COMPUTER AND INFORMATION SCIENCE

Goals: parsimonious modeling

- Scope: methods research for data analysis
- Learning parsimonious models from large and high-dimensional noisy data sets
- Parsimonious = compact, sparse, economical
- *Mission: model representation* should be as simple as possible or minimal
- Computationally intensive algorithms
- Good basis for interpretation


Jaakko Hollmén: Parsimonious modeling research group 3


 HELSINKI UNIVERSITY OF TECHNOLOGY
 LABORATORY OF COMPUTER AND INFORMATION SCIENCE

Collaborative projects

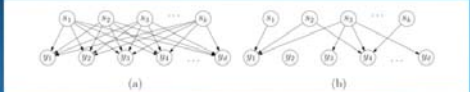
- Identification of diagnostic markers for asbestos-exposure related lung cancer
- Analysis of dependencies in environmental time series
- Multiple data sources in functional genomics for improving genome-wide inferences
- (Analysis of road accidents and road conditions)

Jaakko Hollmén: Parsimonious modeling research group 4


 HELSINKI UNIVERSITY OF TECHNOLOGY
 LABORATORY OF COMPUTER AND INFORMATION SCIENCE

Metaphor: diagnostic model


- Latent diseases and observed symptoms
- Abstract away the irrelevant!



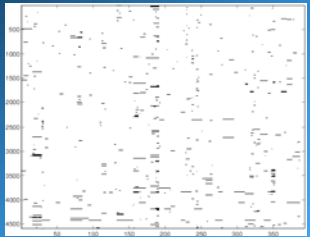
(a) (b)

- Input selection, sparse regression, factor analysis, PCA, ICA, mixture modeling, subspace clustering

Jaakko Hollmén: Parsimonious modeling research group 5


 HELSINKI UNIVERSITY OF TECHNOLOGY
 LABORATORY OF COMPUTER AND INFORMATION SCIENCE

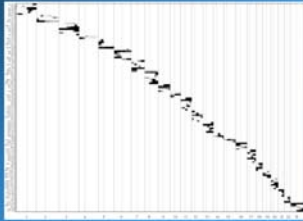
DNA copy number amplification



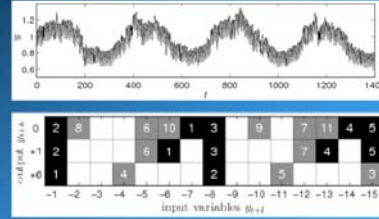
Jaakko Hollmén: Parsimonious modeling research group 6



Amplification patterns for cancer



Sequential Input Selection Algorithm



Past, present and future

- Time series analysis and prediction methods
- Stream mining
- Exploratory analysis
- Ecological applications
- Bioinformatics applications



Parsimonious modeling

“Simple models from complex phenomena”