The Mathematics of Big Data

Joerg Fliege

Professor for Operational Research

Head of Operational Research Group

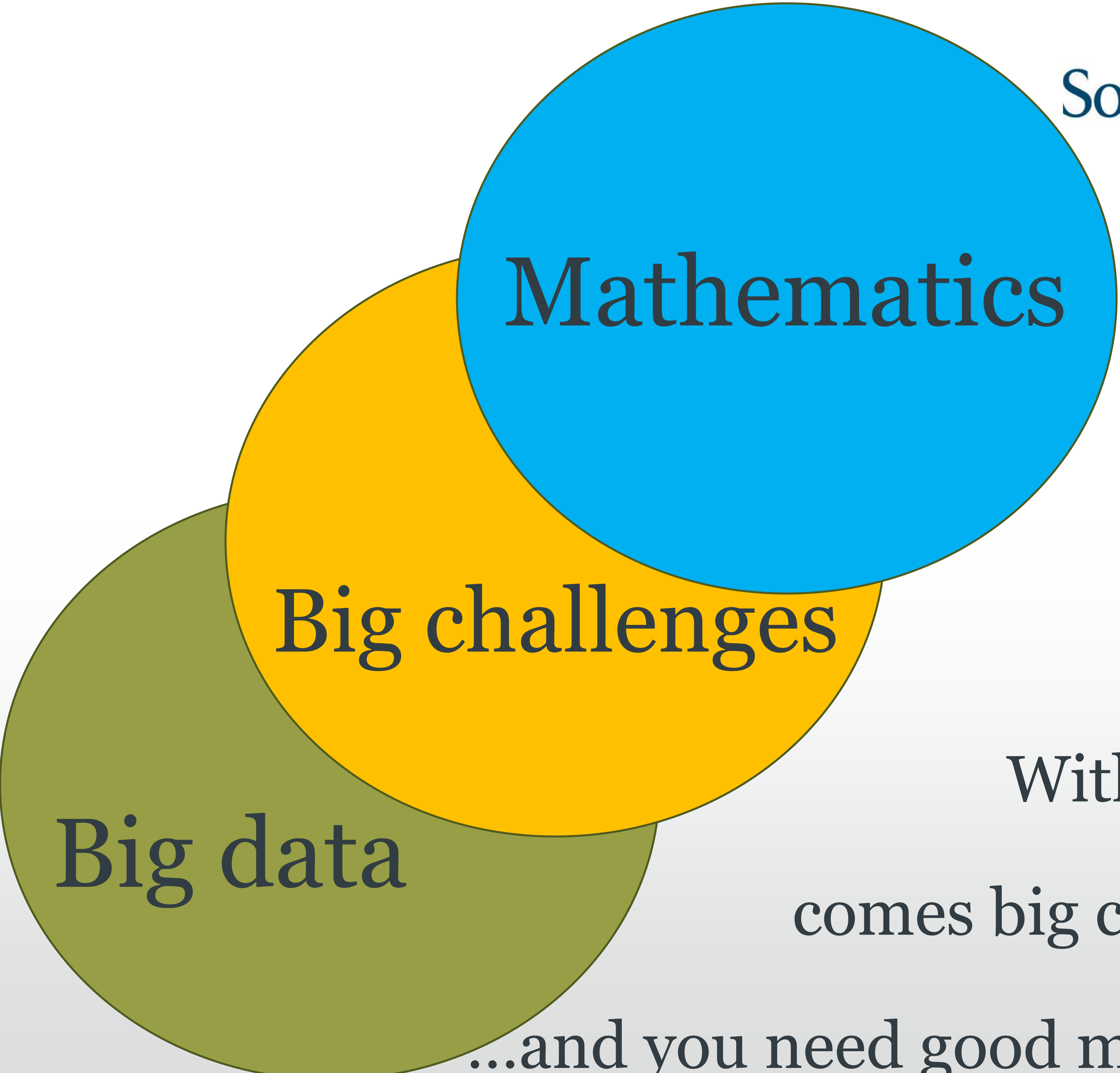
Department of Mathematical Sciences

University of Southampton

Big data

UNIVERSITY OF Southampton

UNIVERSITY OF Southampton



With big data...

comes big challenges...

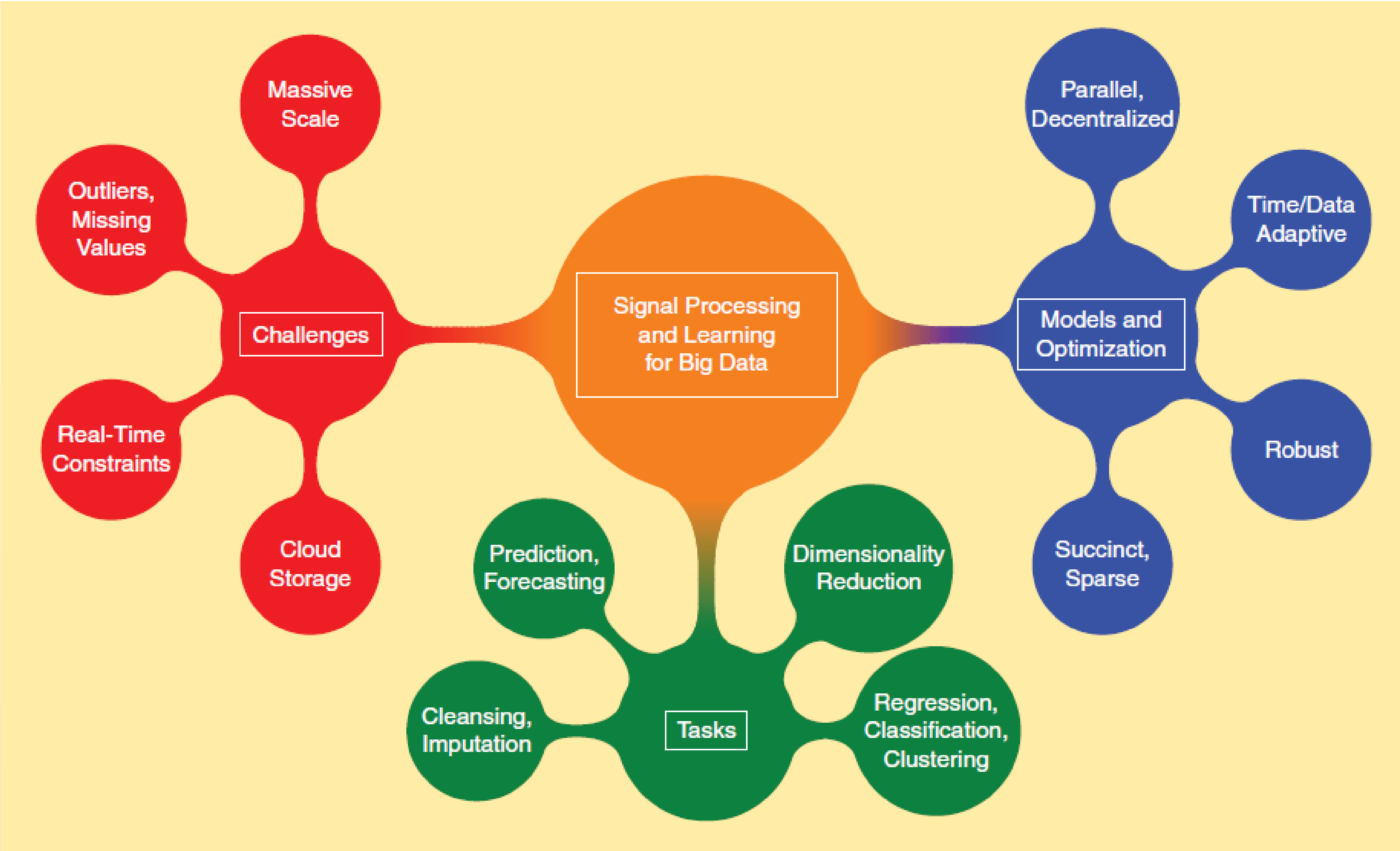
...and you need good mathematics

Big data: challenges

UNIVERSITY OF Southampton

- Central processing infeasible
- Central storage infeasible
- Streaming data: real-time learning
- Streaming: no revisiting of past entries
- Need to revisit old tools from signal processing and statistical learning

Big data: challenges, tasks, and optimization



Example

Here is your data:

Example

Here is your data:

$$Y \in \mathbb{R}^{N \times T}$$

Example

UNIVERSITY OF Southampton

Here is your data:

$$Y \in \mathbb{R}^{N \times T}$$

with N or T huuuuuuuuuuuge.

Example

Here is your data:

$$Y \in \mathbb{R}^{N \times T}$$

with N or T huuuuuuuuuuuge.

(Ex.: traffic data, N traffic links, T time slots.)

Example

$$Y \in \mathbb{R}^{N \times T}$$

We want to decompose Y into

"background data" / trend $L \in \mathbb{R}^{N \times T}$

with L low rank matrix

Example

$$Y \in \mathbb{R}^{N \times T}$$

We want to decompose Y into

"background data" / trend $L \in \mathbb{R}^{N \times T}$

with L low rank matrix

"patterns/clusters/outliers" $S \in \mathbb{R}^{M \times T}$

with S sparse

Example

$$Y \in \mathbb{R}^{N \times T}$$

We want to decompose Y into

"background data" / trend $L \in \mathbb{R}^{N \times T}$

with L low rank matrix

"patterns/clusters/outliers" $S \in \mathbb{R}^{M \times T}$

with S sparse

& modelling/measurement errors $V \in \mathbb{R}^{N \times T}$

Solve

$$Y \approx L + DS + V$$

with some "dictionary matrix" D.

But not all entries of Y are important, so use a projection operator and solve

$$\mathcal{P}(Y) \approx \mathcal{P}(L + DS + V)$$

But how do we model L low rank and S sparse?

Write the task as an optimisation problem:

$$\min_{L,S} \|\mathcal{P}(Y - L - DS)\|_F + \lambda\|L\|_* + \omega\|S\|_0$$

Weight λ controls rank penalty.

Weight ω controls sparsity penalty.

Consider

$$\min_{L,S} \|\mathcal{P}(Y - L - DS)\|_F + \lambda \|L\|_* + \omega \|S\|_0$$

Weight λ controls rank penalty.

Weight ω controls sparsity penalty.

One rich, versatile model that explains data parsimoniously and succinctly.

$$\min_{L,S} \|\mathcal{P}(Y - L - DS)\|_F + \lambda\|L\|_* + \omega\|S\|_0$$

This approach subsumes

- Principle component analysis, robust PCA
- Dictionary learning
- Compressed sampling, compressed sensing
- Subspace clustering
- Nonnegative matrix factorization
- Missing value imputation
- Regression
- Kernel-based learning
- Dimensionality reduction

$$\min_{L,S} \|\mathcal{P}(Y - L - DS)\|_F + \lambda\|L\|_* + \omega\|S\|_0$$

This approach subsumes

- Principle component analysis, robust PCA
- Dictionary learning
- Compressed sampling, compressed sensing
- Subspace clustering
- Nonnegative matrix factorization
- Missing value imputation
- Regression
- Kernel-based learning
- Dimensionality reduction

One mathematical model to rule ~~them all~~
a lot of other approaches

Algorithms

UNIVERSITY OF Southampton

$$\min_{L,S} \|\mathcal{P}(Y - L - DS)\|_F + \lambda\|L\|_* + \omega\|S\|_0$$

- ADMM: alternating direction method of multipliers
- DR: Douglas-Rachford algorithm
- BCDM: block-coordinate descent methods
- K-SVD
- Mardani-Mateos-Giannakis
- Iterative subgradient

Algorithms

$$\min_{L,S} \|\mathcal{P}(Y - L - DS)\|_F + \lambda\|L\|_* + \omega\|S\|_0$$

- ADMM: alternating direction method of multipliers
- DR: Douglas-Rachford algorithm
- BCDM: block-coordinate descent methods
- K-SVD
- Mardani-Mateos-Giannakis
- Iterative subgradient

Decentralized

Parallelizable

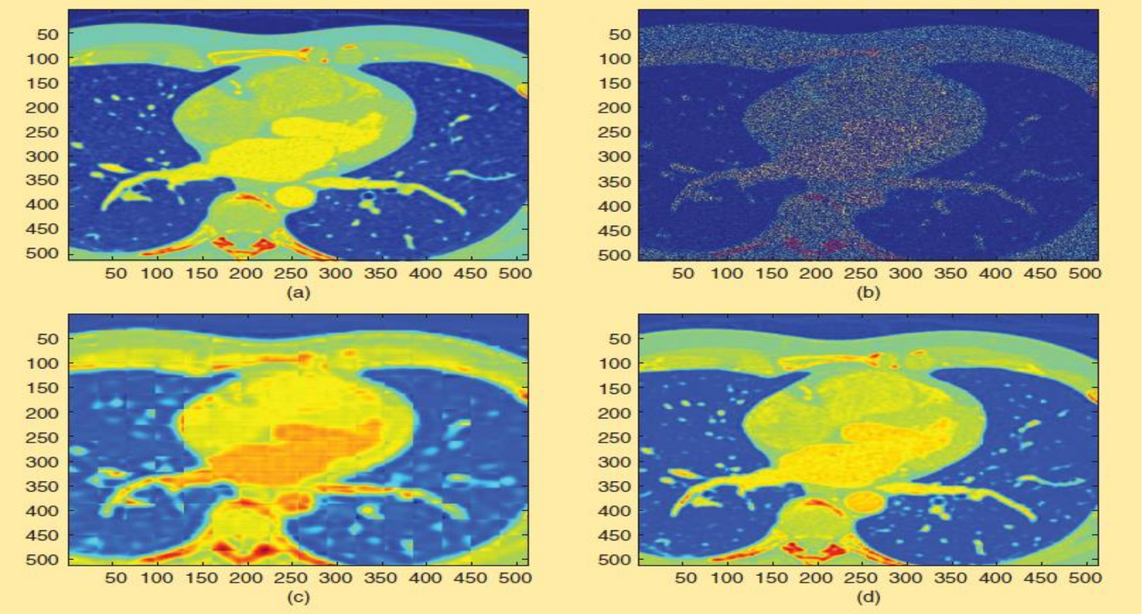
Robust

Online

Scalable

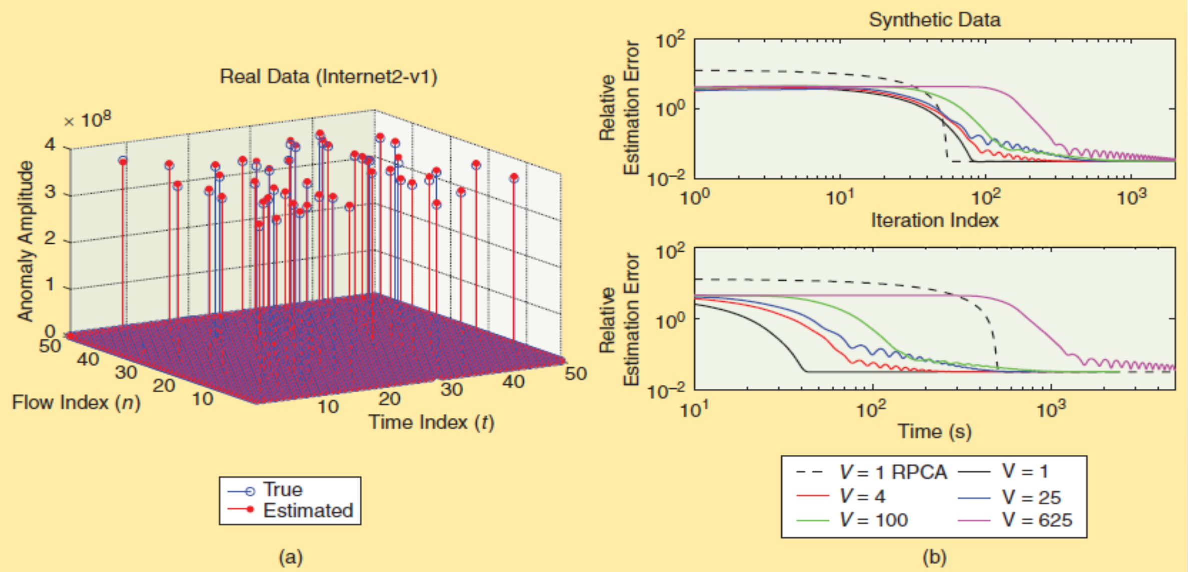
Convergence guarantee: we know they always work!

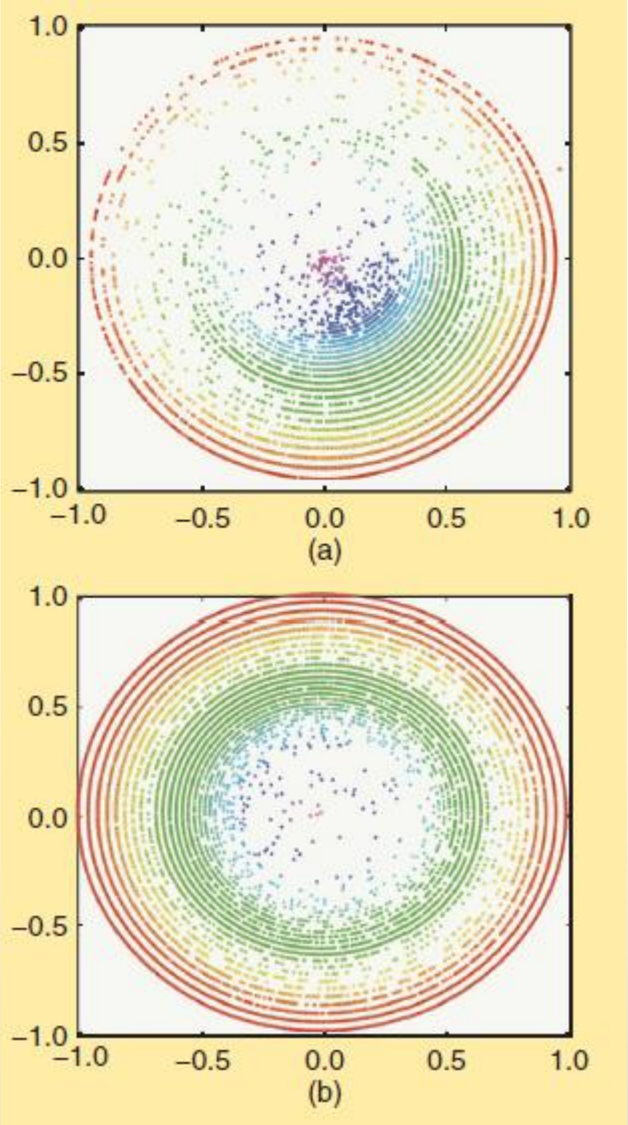
Applications



Imputation of MRI data

Traffic: outlier detection & analysis

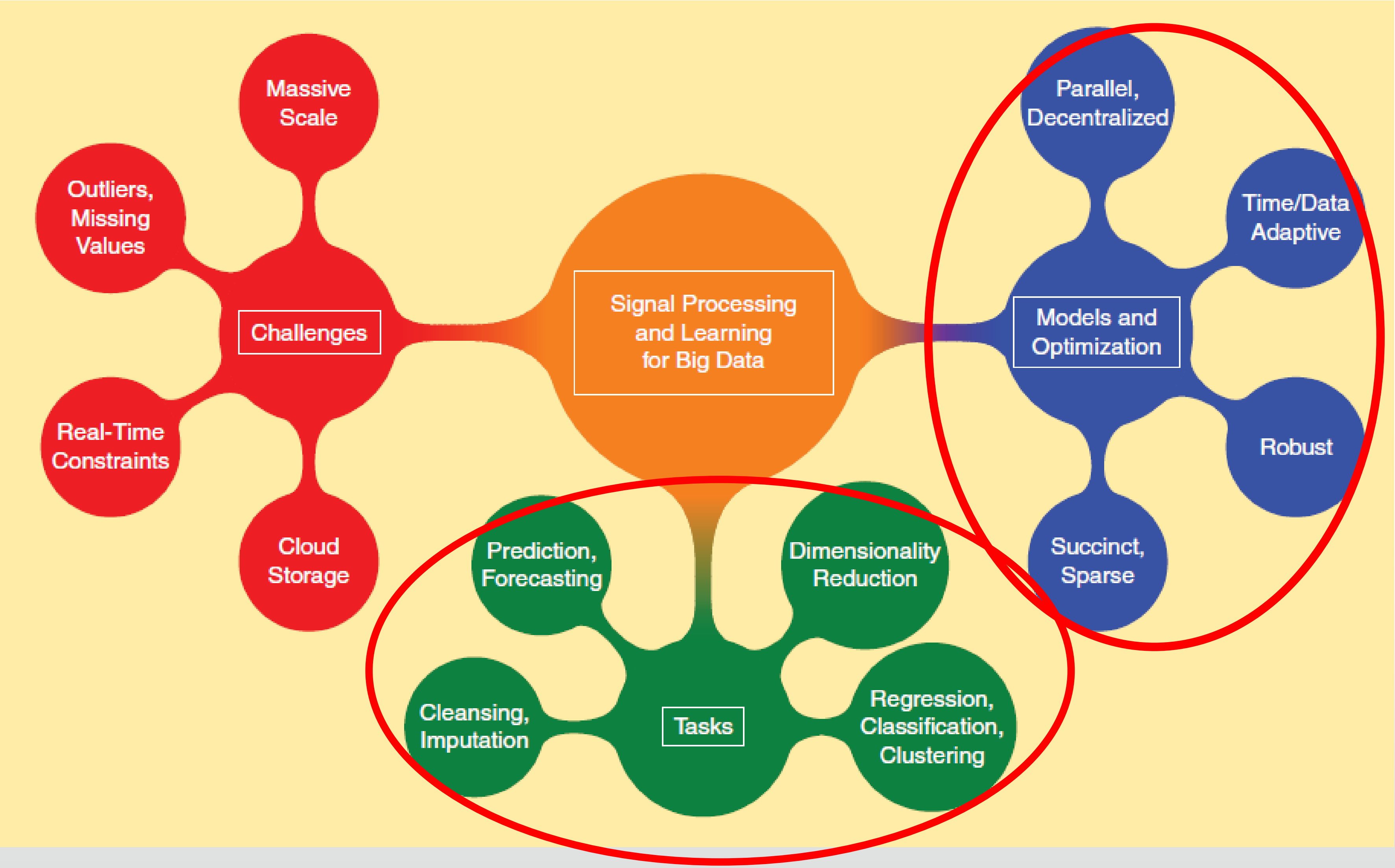




Dynamic network visualization

Conclusions

UNIVERSITY OF Southampton



Mathematics: we are here to help.