Missing Value Estimation for Hierarchical Time Series: A Study of Hierarchical Web Traffic

Zitao Liu

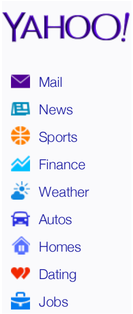
University of Pittsburgh

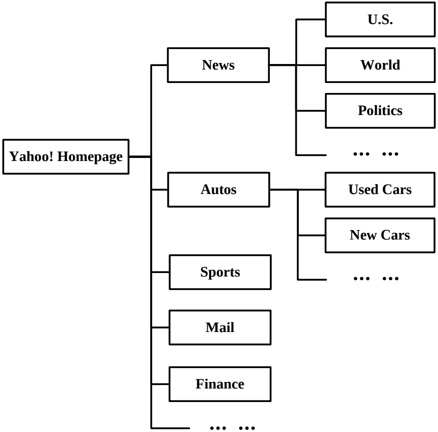
ztliu@cs.pitt.edu

November 16, 2015

† This is the joint work with Chris Yan, Jimmy Yang and Milos Hauskrecht.

Hierarchical Time Series

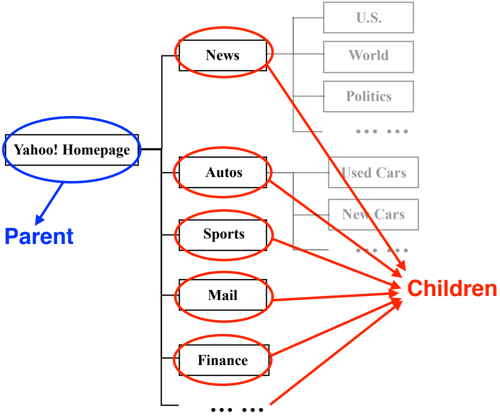




What is **hierarchical time series**(HTS)? Yahoo! web pages are arranged in certain hierarchy and their **daily page views** become a hierarchical time series.

Hierarchical Time Series

Time series are organized in a hierarchical tree structure and they are **consistent between hierarchy levels**.



Consistent: Parent = Child 1 + Child 2 + Child 3 + ... + Child n

Why we care about modeling HTS ?

- Resource management.
- User behaviors understanding.
- Advertisement pricing policy.

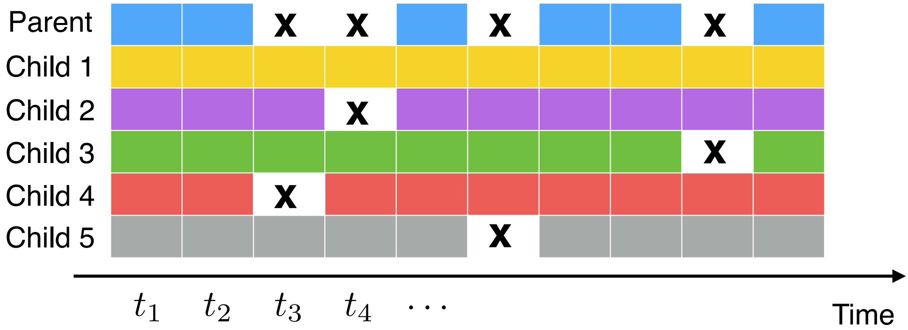
Problem & Goal

However, missing values occur:

- machine failures
- networking disturbances
- human mistakes

Missing values will contaminate other time series through the hierarchy consequentially.

Problem



Accurately estimate the missing values.

s.t

Estimation is hierarchically consistent.

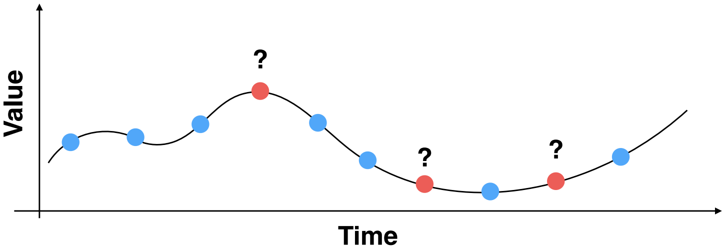
HTSImpute

In this work, we develop a new missing value estimation algorithm *HTSImpute* which

- utilizes the temporal dependence information within each individual time series (**LOcal regrESSion (LOESS)**)
- exploits the intra-relations between different time series (**Subspace Projection**)
- guarantees hierarchical consistency (**Hierarchical Consistency Projection**)

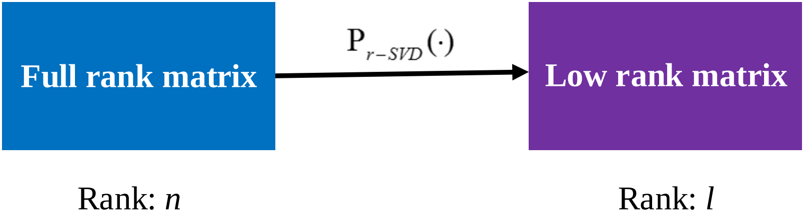
HTSImpute - LOESS

Use LOESS to initially estimate the missing values.

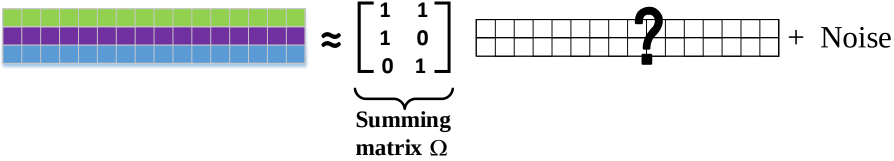


Advantages:

- nonparametric
- robust
- locally weighted


Full rank matrix
$P_{r-SVD}(\cdot)$
Low rank matrix
Rank: n
Rank: l

HTSImpute - Hierarchical Consistency Projection



$$\mathbf{Y} = \Omega \cdot \hat{\mathbf{L}} + \epsilon_{\mathbf{Y}}$$

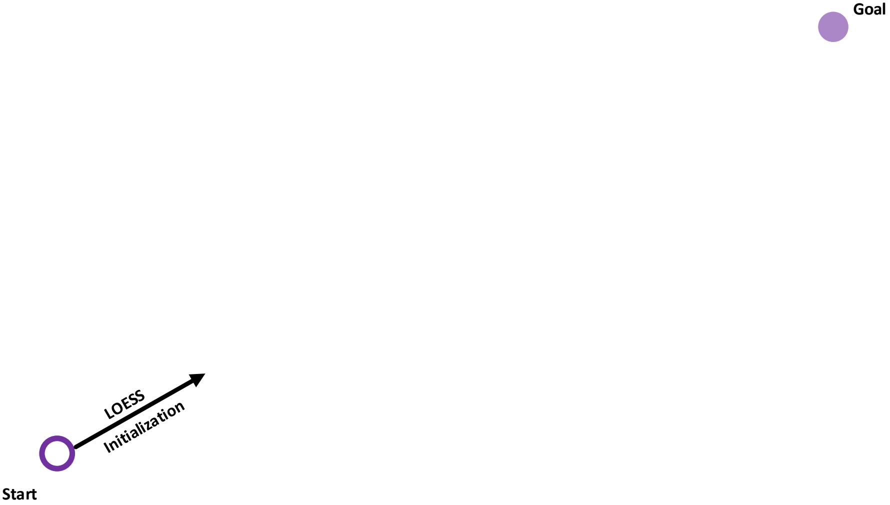
where $\hat{\mathbf{L}}$ is the "true" estimate of all leaf time series.

We define the hierarchical consistency projection operator using ordinary least square as follows:

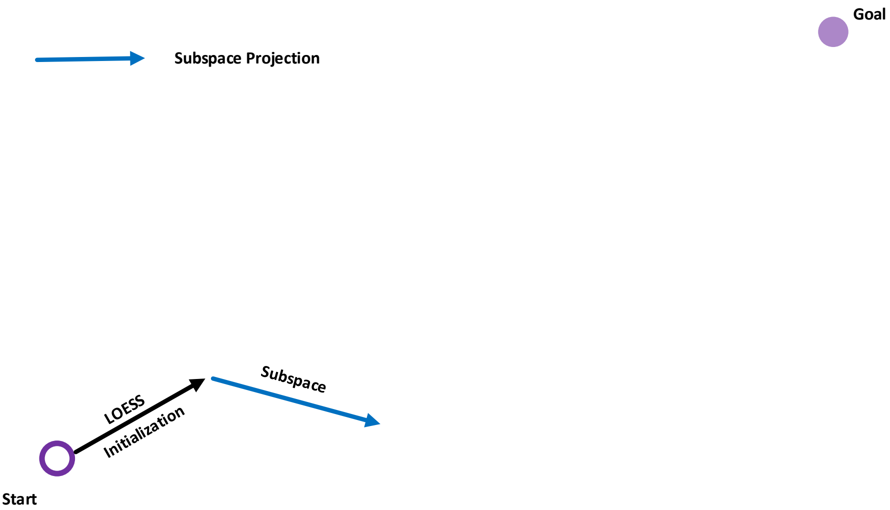
$$\mathcal{P}_{\mathsf{HTS}}(\mathbf{Y}, \boldsymbol{\Omega}) = \boldsymbol{\Omega}\hat{\mathbf{L}} = \boldsymbol{\Omega}(\boldsymbol{\Omega}^\top \boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^\top \mathbf{Y}$$

HTSImpute - Idea Illustration

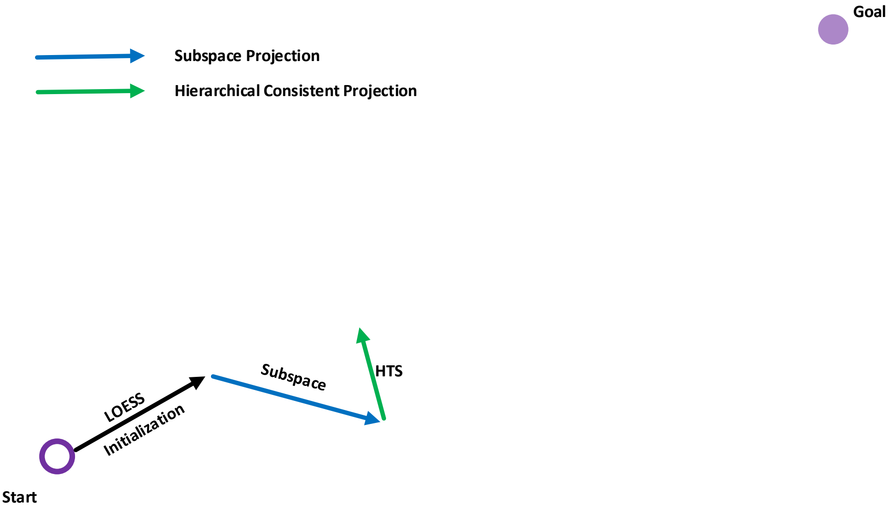
HTSImpute - Idea Illustration



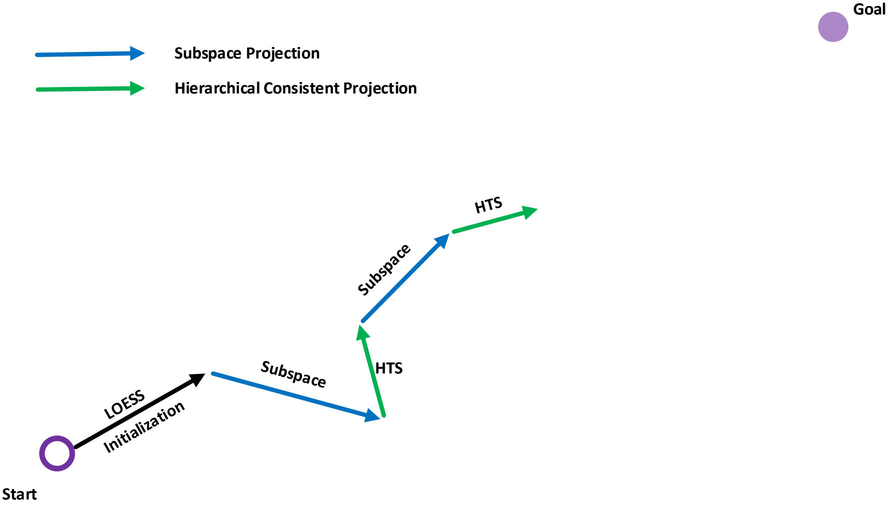
HTSImpute - Idea Illustration



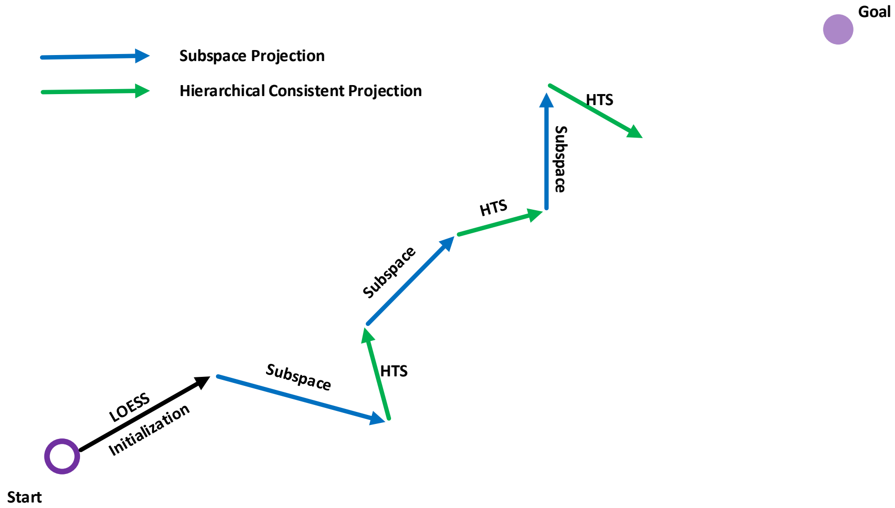
HTSImpute - Idea Illustration



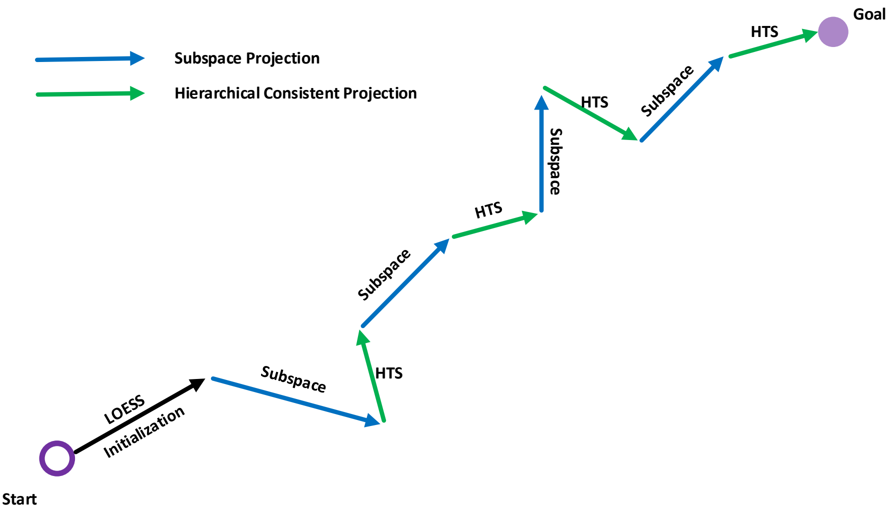
HTSImpute - Idea Illustration



HTSImpute - Idea Illustration



HTSImpute - Idea Illustration



Experiments - Dataset

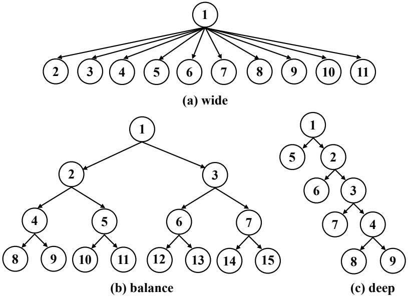


Figure 1: Synthetic data.

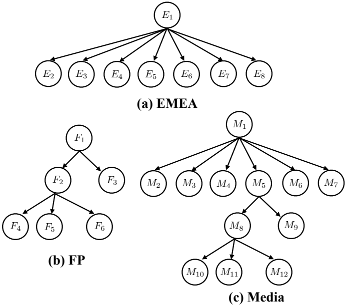


Figure 2: Yahoo! web traffic data.

Experiments - Metrics

Avg-MAPE: measures the estimation accuracy.

$$``\frac{\text{Estimated Value} - \text{True Value}}{\text{True Value}}"$$

Avg-HCG: measures the hierarchical consistency.

"Estimated Parent Value − Sum of Estimated Child Values"

Experiments - Baseline

- Regression Methods
  - Local regression (LOESS)
- Subspace Methods
  - Matrix Factorization (MF)
  - Matrix Completion (MC) using softImpute
  - weight Low Rank Approximation (wLRA)
- Latent Variable Models
  - probabilistic PCA (pPCA)

Experiments - Results

Table 1: Avg-MAPE results on *FP* dataset.

# MP (%)	1	3	5	10	15	20
LOESS	11.57	11.84	11.66	11.78	11.82	11.78
NMF_KL	15.12	14.11	13.83	14.48	14.74	14.61
NMF_Euclidean	21.19	17.75	16.49	17.41	15.70	15.97
MC	57.30	44.27	46.76	57.11	63.37	66.85
pPCA	101.51	100.17	100.21	100.03	100.01	100.19
wLRA	42.39	73.64	96.91	41.54	34.03	30.10
HTSImpute	**7.19**	**7.48**	**7.40**	**7.87**	**8.22**	**8.36**

Table 2: Avg-HCG results on *FP* dataset ($\log_{10}$ scale).

# MP (%)	1	3	5	10	15	20
LOESS	-2.61	-2.14	-1.93	-1.67	-1.54	-1.45
NMF_KL	-2.52	-2.04	-1.95	-1.71	-1.60	-1.51
NMF_Euclidean	-2.01	-1.88	-1.80	-1.62	-1.55	-1.47
MC	-1.91	-1.35	-0.96	-0.19	0.30	1.34
pPCA	0.79	0.97	1.39	1.49	2.18	1.78
wLRA	-5.65	-4.83	-1.58	-1.67	-1.55	-1.47
HTSImpute	**-16.66**	**-16.32**	**-16.04**	**-15.80**	**-15.71**	**-15.61**

Conclusion

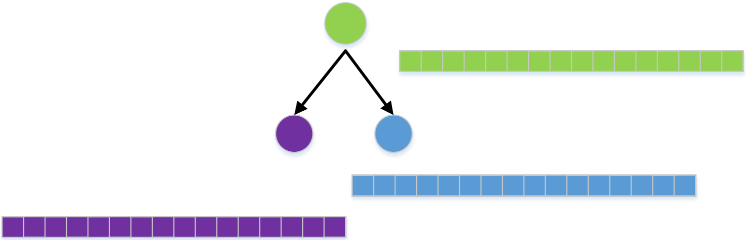
In this work, we have presented a algorithm for HTS missing value estimation, specializing in

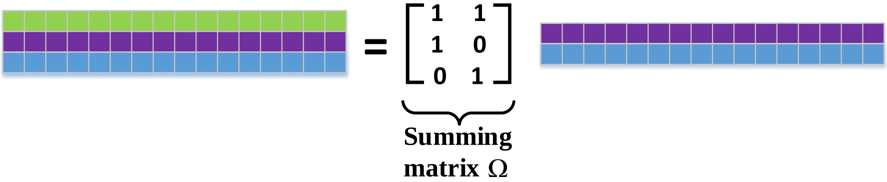
- taking advantage of temporal dependence information within each individual time series.
- utilizing intra-relations between different time series across the hierarchy.
- providing high satisfaction of the hierarchical consistency.

Thank you

November 16, 2015

HTSImpute - Hierarchical Consistency Projection





HTSImpute - Hierarchical Consistency Projection