

Spatio Temporal Neural Network for missing network data imputation

Le Tien Thanh – Le Van An – Yusheng Ji

Motivation

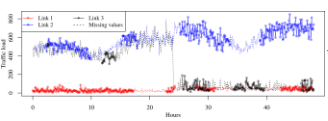
Missing values appear in most multivariate time series, especially in the monitored network traffic data due to high measurement cost, or unexpected faults. In the networking field, the problem of missing data prevents advanced analysis and downgrades the performance of downstream applications such as traffic engineering. Therefore, we are conducting research to impute the missing data with minimum error by applying deep learning technique.

Preliminaries

- What is network measurement data?
- Why do we need data imputation?
- What are previous data recovery methods?
- What is the solution approach?

Research content

- Fig1. Missing network traffic data
- Time series with missing values
 - L time steps
 - D flows/links



- Temporal Decomposition Module to encode the feature of observed data on time dimension.
 - Spatial Decomposition Module to encode the feature of each network flows/links.
 - Combine the features and predict \hat{X} that best match X .
- ⇒ We named it Convolutional Recurrent Imputation of Network Data - **CRIND**

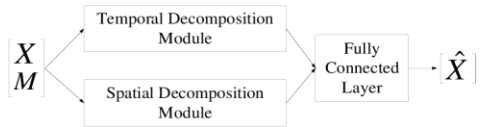


Fig2. Model overview

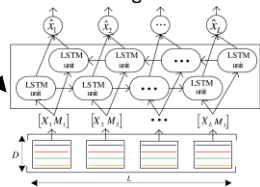


Fig3. A Bidirectional Long-short term memory (LSTM) - learn how different time steps related

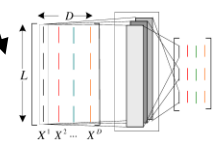


Fig4. 1D convolutional – learn how different links/flows related

Experimental studies



Fig5. The Abilene Network

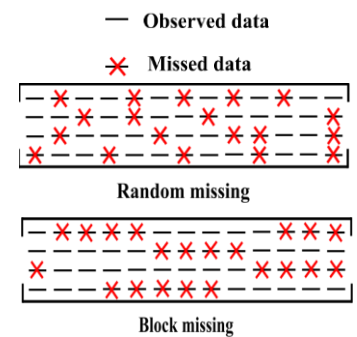


Fig6. Missing scenarios

Dataset:

- Abilene_TM: traffic between two nodes
 - Abilene_LL: traffic between two adjacent node
- Two missing scenarios: random and block
Missing rate: 10% to 90%

