BRIGHT - Drift-Aware Demand Predictions for Taxi Networks

Amal Saadallah[1], Luís Moreira-Matias[2], Ricardo Sousa[3], Jihed Khiari[4], Erik Jenelius[5], João Gama [3]

[1]Chair of Artificial Intelligence, TU Dortmund, Germany ,[2]Kreditech Holding SSL, Hamburg, Germany
[3]LIAAD/INESC TEC, University of Porto, Portugal, [4]Data Science Applications Group, NEC Laboratories Europe, Germany
[5]Dep. of Civil and Architectural Engineering, KTH Royal Institute of Technology, Sweden

Abstract

- Matching taxi supply with demand is one of the biggest challenges in taxi industry due to the dynamic behavior of urban mobility patterns.
- Building a data-driven real-time taxi-dispatching recommender system is a promising solution especially with the increasing availability of massive broadcast GPS data.
- Existing systems are based on strong assumptions such as stationary demand distributions and finite training sets, which make them inadequate for modeling the dynamic nature of the network.
- **BRIGHT** framework is proposed to solve the aforementioned issues: **BRIGHT** is a drift-aware supervised learning framework which aims to provide accurate predictions for short-term horizon taxi demand quantities through a creative ensemble of time series analysis methods that handle distinct types of concept drift.

Problem formulation

The taxi passenger demand prediction problem:

- *Given:* Historical realizations of random demand processes $Y^i, \forall i \in \mathbb{K}$ and $\mathbb{Y}$ denoting the following Euclidean space:

$$\mathbb{Y} = \begin{pmatrix} y_1^1 & y_2^1 & \cdots & y_{t-1}^1 & y_t^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_1^i & \cdots & \cdots & y_{t-1}^i & y_t^i \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_1^K & \cdots & \cdots & y_{t-1}^K & y_t^K \end{pmatrix} = \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_{t-j} \\ \vdots \\ \mathbf{Y}_t \end{pmatrix}^T$$

- *Objective:* Forecast all the K realizations of $\mathbb{Y}$ at instant $t+1$.
- *Challenges:* Non-stationary random processes, presence of concept-drifts,...

Technical Details

BRIGHT operates in three different blocks:

1. Baseline learners are trained to produce demand forecasts.
2. For multivariate forecasting models : the random processes used to model the service counts value for a given ROI (i.e Region Of Interest) can be historical values from some other ROIs denoted here as *neighborhood*, or exogenous variables (e.g. weather-related). The *neighborhood* selection is made in a drift-aware fashion.
3. Baseline models are combined by an online meta-learner in an ensemble fashion
   - Base learners are grouped offline using their historical outputs (e.g. using a Gaussian Mixture Model learned through EM algorithm) into at least two families.
   - An online model selection procedure based on Page Hinckley test (PH), is triggered where only one method per group (winner) is exposed to the final meta-learning process.
   - Winners output are combined in a weighted average where the weights are inversely proportional to their recent loss.

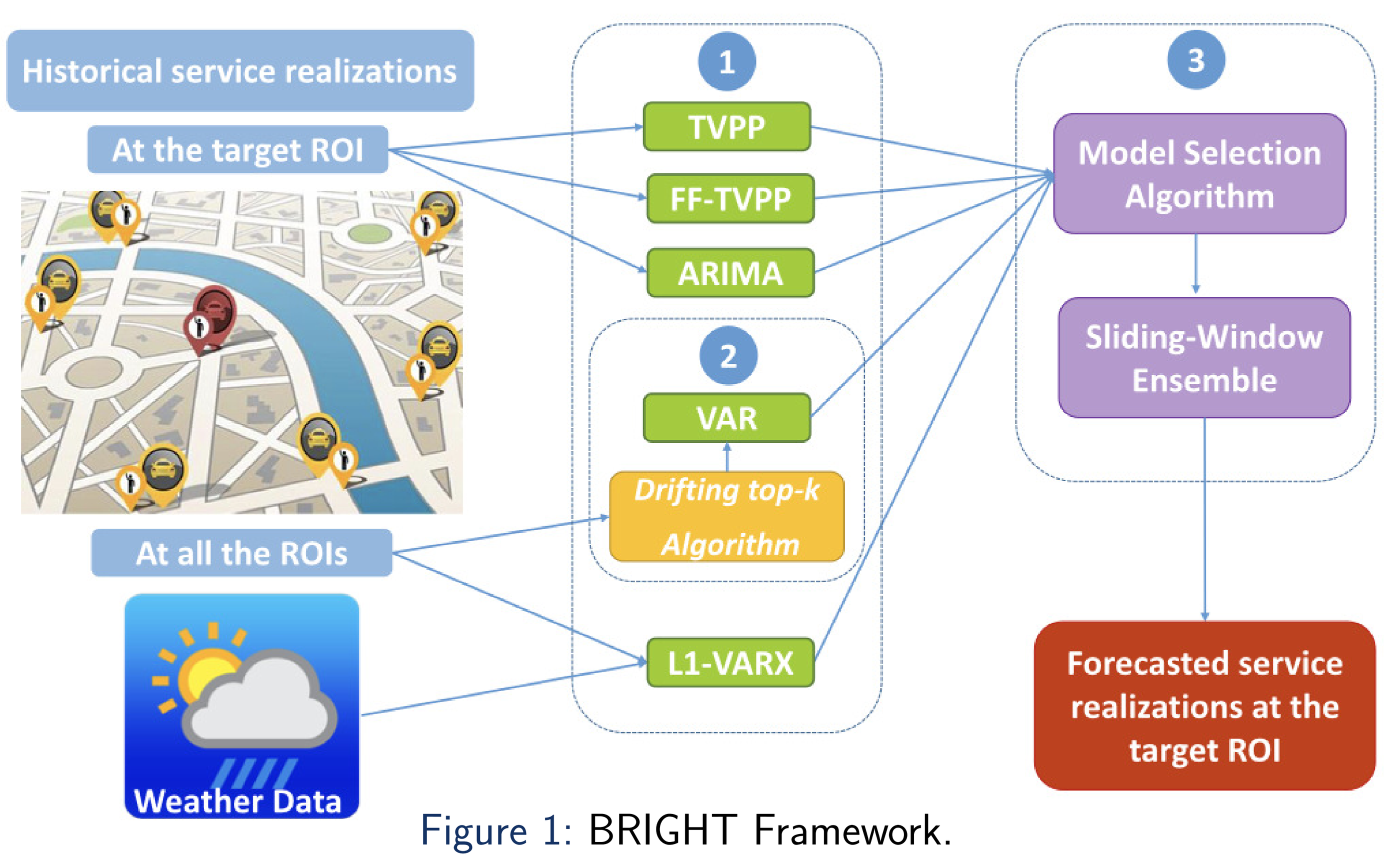
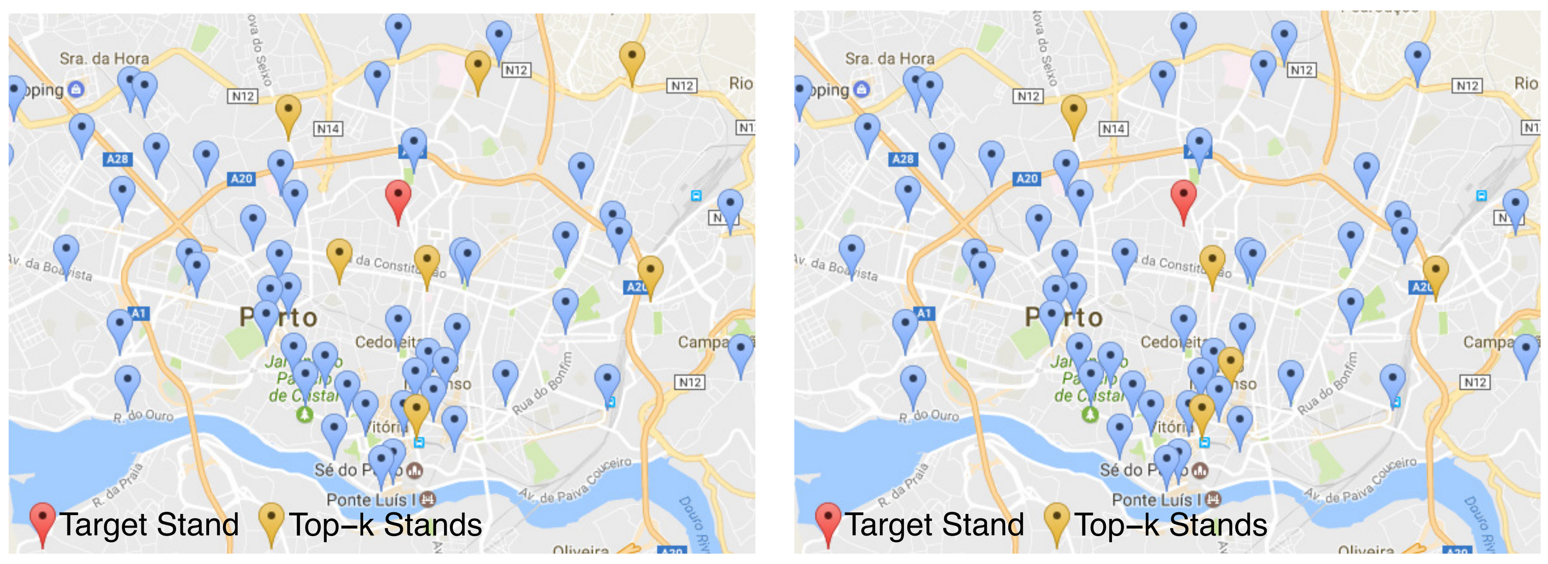


Figure 1: BRIGHT Framework.

Drift-Aware VAR with top-k Selection

- To realize a drift-aware regularization of a VAR model, a two-stage online algorithm denoted `Drifting Top-k` is devised to select a neighbourhood $\mathbb{L}$ for each target variable in every random process $Y^i, \forall i \in \mathbb{K}$ of k other processes ($|\mathbb{L}| = k$ is a user-defined hyperparameter).
- This neighborhood is selected through an incremental similarity measure by the closest k processes to the target.
- The distance between the two most dissimilar random processes within the same neighbourhood sets its *boundary* under a form of a logical *diameter*. If this *boundary* diverges *significantly* over time, a drift is assumed to be in place. This is incrementally verified using the well-known Hoeffding bound.
- Weather is known to have a huge impact on mobility patterns. In order to add this knowledge to BRIGHT, multiple time series of meteorological variables are incorporated to train a VAR model with exogenous explanatory variables(VARX). L1-regularization is also applied during the training.



(a) 00:00am-08:00am (b) 11:30am-05:30pm

Figure 2: Illustration of the dynamic neighborhood setting of the drift-aware VAR model in a ROI from Porto taxi dataset.

Experimental Results

- Experiments were conducted using three large-scale real-world transportation networks in Porto (Portugal), Shanghai (China) and Stockholm (Sweden) and a synthetic data where multiple distinct drifts were artificially induced.
- Fig.4 shows an example of an incremental drift:
  1. The first PH alarm was triggered before the loss increase of the current *family winner*:VAR.
  2. A *Hoeffding* alarm is triggered later - flagging that the current neighbourhood structure of this ROI changed due to the occurrence of a drift. Consequently, its neighbourhood is recomputed, the VAR model order/parameters are updated and the loss decays for its normal range.
  3. During the adaptation period, a fairly simpler model (ARIMA) was put in place (i.e. *winner*) to minimize the loss increase.

⇒ **This informed adaptation ability is a key point of BRIGHT.**

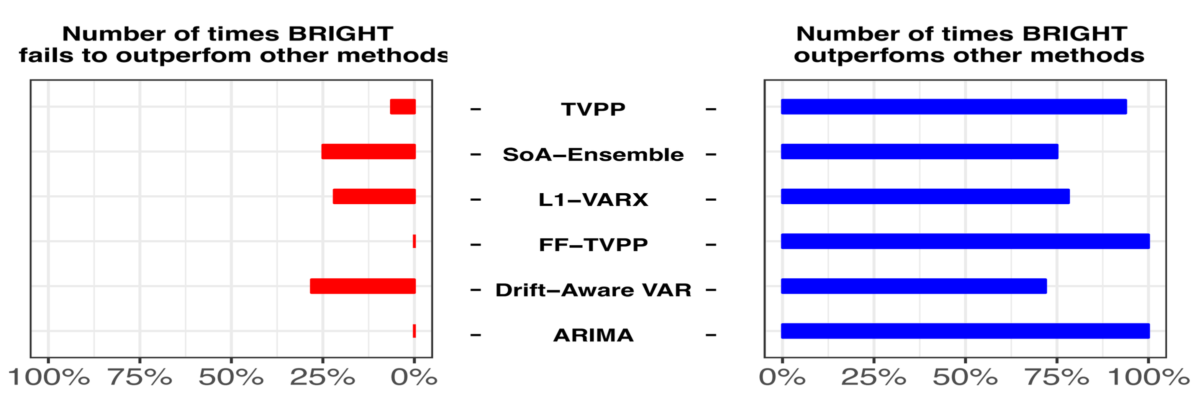


Figure 3: Wilcoxon signed-rank test results using pairwise comparisons on a *BRIGHT-against-all*.

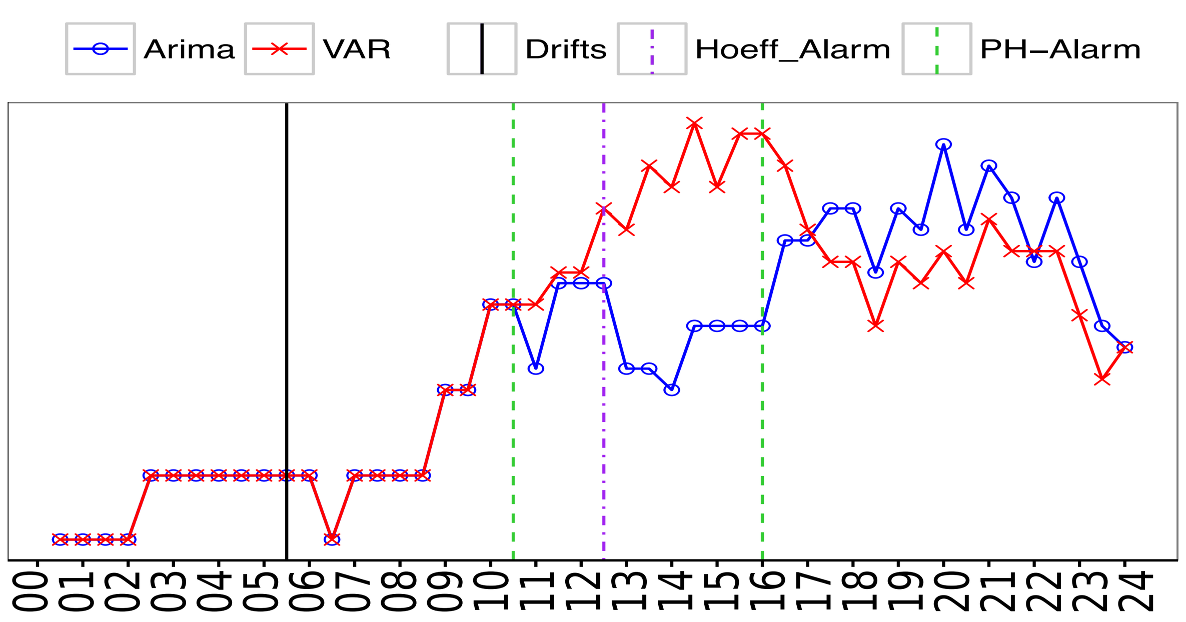


Figure 4: MSE of two methods from the same *family* for 24h on the synthetic data

References

- Amal Saadallah, Luis Moreira-Matias, Ricardo Sousa, Jihed Khiari, Erik Jenelius, and Joao Gama. Bright-drift-aware demand predictions for taxi networks. IEEE Transactions on Knowledge and Data Engineering, 2018.