Learning Temporal Rules to Forecast Instability in Intensive Care Patients

Mathieu Guillame-Bert

Postdoctoral Fellow at Carnegie Mellon University (mathieug@andrew.cmu.edu)

with Artur Dubrawski, Lujie Chen, Andre Holder, Gilles Clermont, Marilyn Hravnak, and Michael R. Pinsky

This work has been partially funded by the National Science Foundation (awards 0911032, 1320347) and the National Institutes of Health (R01NR013912)

- Cardio-respiratory instability can be life threatening in critically ill patients.
- Monitoring devices which track single vital sign (VS) signals independently to observe VS deviations and detect instability are problematic
  - True instability is rarely a single-parameter abnormality
  - Unstable patients can deteriorate quickly → need to rapidly (yet reliably) detect onset of instability
  - **Forecasting** is more useful but more difficult than **detection**
- **Hypothesis**
  It is best to simultaneously use multiple vital signs, even though models tend to get complex (hard to interpret) when multivariate data is sparse and noisy.
- **Our goal**
  Learn human-interpretable multivariate models to forecast instability.

Methodology

1

- Acquire vital signs data
  HR, RR, SpO_2 and Blood pressure

2

- Extract a large set of secondary variables (features)
  Statistical analysis, quality of signal, etc.

3

- Learn a set of temporal models with the Temporal Interval Tree Association Rule Learner (TITARL) algorithm.
  Temporal association rules.

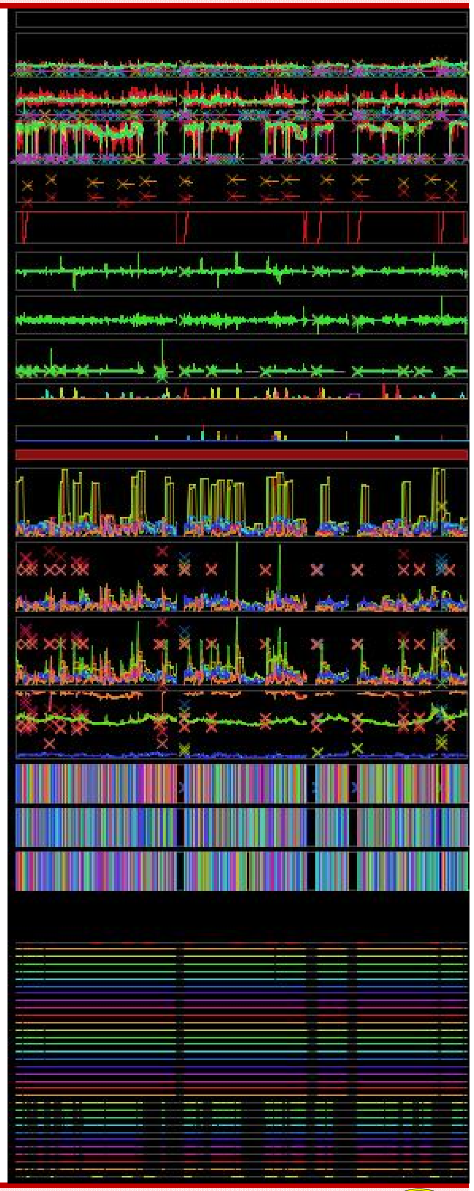
4

- Fuse temporal association rules into a compact ensemble forecasting model

5

- Use the learned model to forecast instabilities
  Report cross-validated results.

Extracting features



Input

The **heart rate**, **respiratory rate**, **SpO_2** level (every 20s) and **blood pressure** (every ~30min) for three patient-years of monitoring data duration. We censored data after the first instability of each patient, leaving 1.43 patient-years of data including 130 episodes of instability.

Processing

A large number of features is extracted from raw data:
Moving averages (uniform, triangular, exponential), median, standard deviation, ranges, derivatives, cumulative sums, hysteresis with various thresholds, quality of signal, MACD derivatives, calendar events (hours, days, etc.).

Output

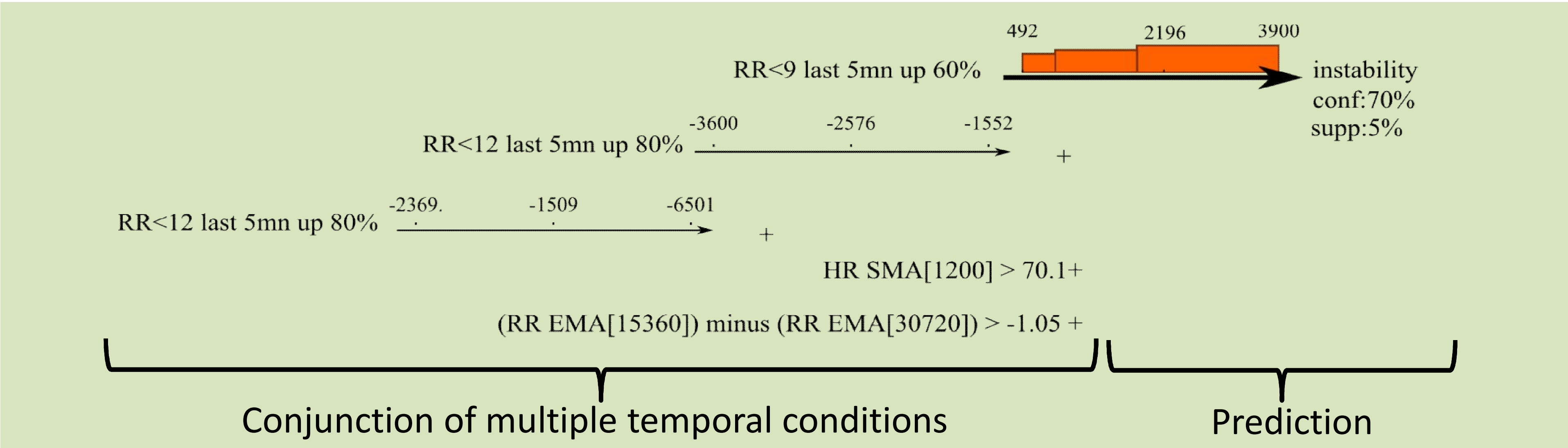
Symbolic and scalar time sequences:

~ 7,400 types of secondary observations (features) per each instance of time and patient

~ 32,300,000 such instances

TITARL (Temporal Interval Tree Association Rule Learner)

- **TITARL** is a data mining algorithm designed to extract temporal association rules from symbolic and scalar event sequences*
- Example rule extracted with TITARL (cross-validated):

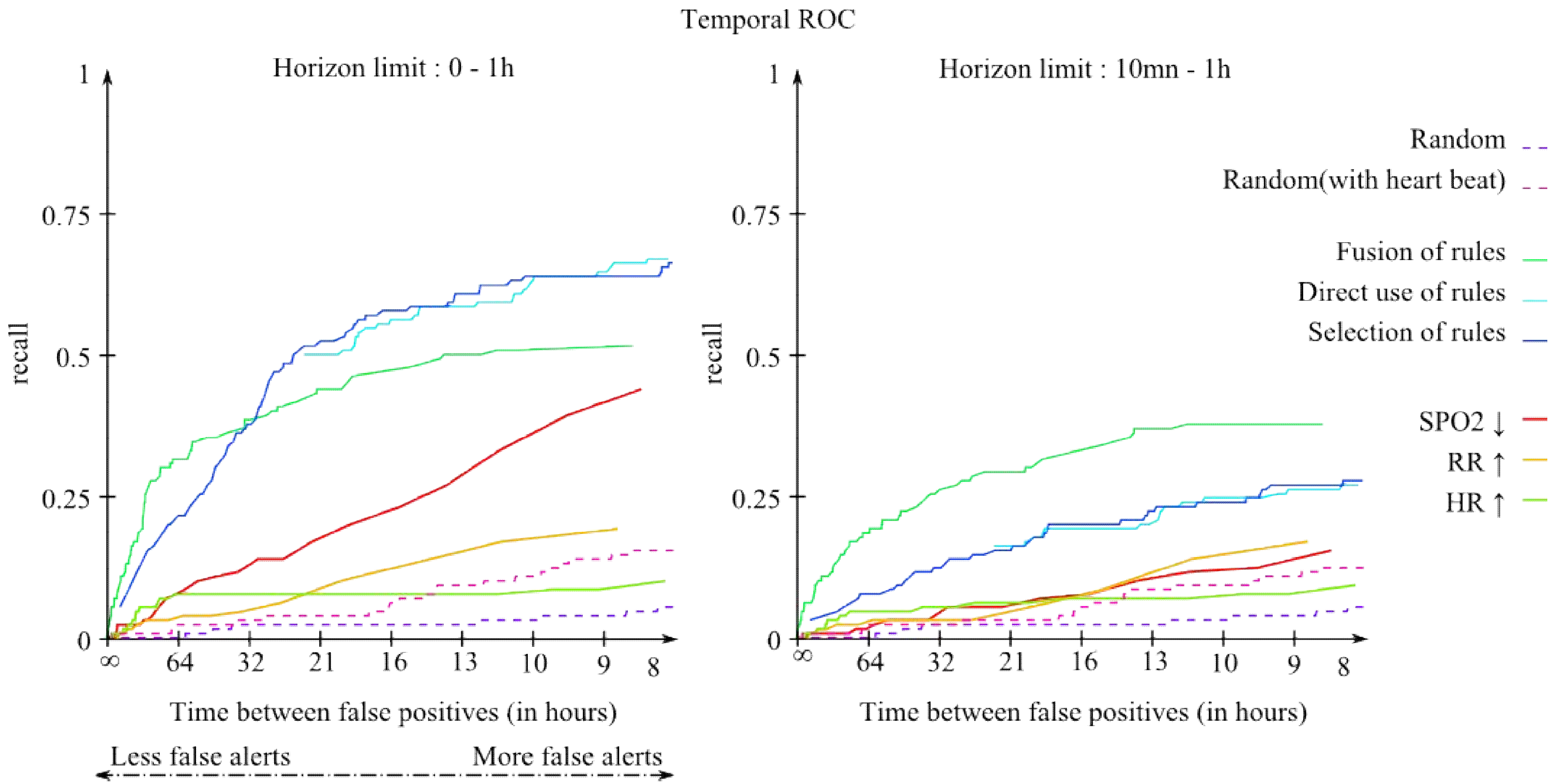


- Example of a condition:

RR<12 for last 5mn up 80% = « RR has been below 12 for more than 80% of the last 5min »

* *New Approach on Temporal Data Mining for Symbolic Time Sequences: Temporal Tree Associate Rules, 2011 IEEE 23rd International Conference on Tools with Artificial Intelligence.*

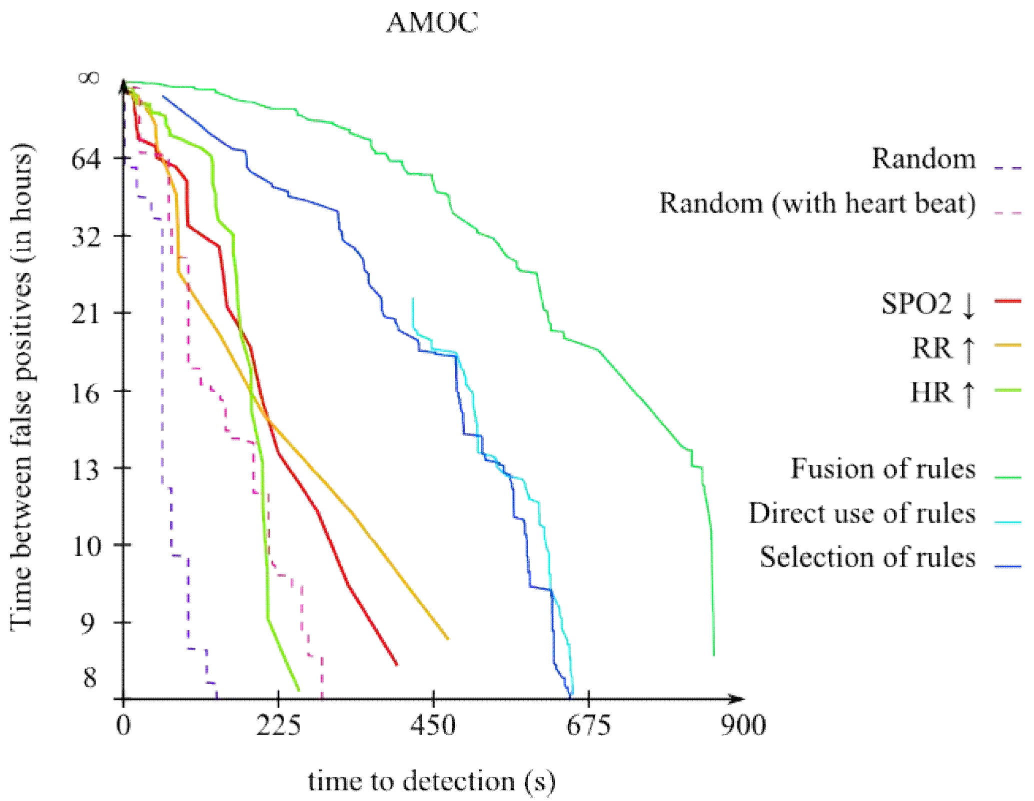
Results 1/2: Forecasting ability



We compare 3 variants of **rule-based forecasting models** (direct use of multiple rules, and two methods of rule fusion) vs. **direct use of vital signs**, and random predictors.

We observed a **multi-fold improvement** of **recall** of the episodes of instability.

Results 2/2: Recall vs. forecasting horizon



The three types of rule based models show comparable recall.

However, **fusion method** (the most advanced technique), shows a **large improvement of the average lead detection time.**

1 false alert every 20 hours

Technique	Recall	Avg. time to detection
Fusion of rules	46%	10mn 38s (638s)
Selection of rules	52%	6m 13s (373s)
Direct use of rules	45%	5m 59s (359s)
Heart rate	8%	2m 44s (164s)
Random (HB)	4%	1m 32s (92s)