GECCO 2015 Industrial Challenge

Farzad Noorian
Computer Engineering Lab
School of Electrical and Information Engineering
The Problem

Supply Temperature Setpoint

Return Temperature

System Supply Temperature

System Power

Boiler
The Data
The missing data

---

<table>
<thead>
<tr>
<th>Data removed in</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>No data removed</td>
<td>2013-11-18 05:12:00 – 2014-01-24 11:02:00</td>
</tr>
<tr>
<td>Supply temperature setpoint</td>
<td>2014-01-24 11:03:00 – 2014-03-24 11:43:00</td>
</tr>
<tr>
<td>Return temperature</td>
<td>2014-05-22 13:25:00 – 2014-07-20 14:05:00</td>
</tr>
<tr>
<td>System power</td>
<td>2014-07-20 14:06:00 – 2014-09-17 14:46:00</td>
</tr>
<tr>
<td>All 4 time series</td>
<td>2014-09-17 14:47:00 – 2014-11-15 14:27:00</td>
</tr>
<tr>
<td>No data removed</td>
<td>2014-11-15 14:28:00 – 2015-01-13 15:08:00</td>
</tr>
</tbody>
</table>
The missing data
The missing data
Machine learning approach with evolutionary model selection
1. Supply setpoint
2. Supply Temperature
Supply Temperature first difference

![Graph showing supply temperature first difference with time series and frequency analysis.]
The reason for choosing \textit{diff} as target

\textbullet\ Ordinary regression:

![Graph showing important in modelling!](image-url)
The other reason for choosing diff as target

› The difference:
The other reason for choosing diff as target

- The difference: Backwards and forwards accumulation merged
The Approach

› Machine learning approach with evolutionary model selection

- Difference Regression
- System Supply Temperature
- Return Temperature
- System Power

Multiclass classification
Supply Temperature Setpoint
Missing all data: forward time-series prediction
Missing all data: backward time-series prediction
Missing all data: Merging backward and forward
Missing all data: Merging backward and forward
The approach

A machine learning pipe-line

Cross-validation

Pre-processing → Feature Extraction → Supervised Learning → Target

- differencing
- logarithm stabilization
- windows of other data

Supply temp. setp. \([t - m, \ldots, t + m]\)
Return temp. \([t - n, \ldots, t + n]\)
Sys. Power \([t - p, \ldots, t + p]\)

\(m, n, p = ?\)

\(C = ?\)
\(\gamma = ?\)

Sys. supply temp. = C

\(\gamma = ?\)
Cross-validation

Training
Training
Training
Testing
Testing
Testing

Parameters

Pre-processing → Feature Extraction → Supervised Learning → Target

SVM Model
Cross-validation

The fitness function!

A Chromosome

Training
Testing
Training
Testing
Training
Testing

Pre-processing → Feature Extraction → Supervised Learning

SVM Model → Target

Parameters

Prediction Error
Evolutionary optimisation: The Grammatical Evolution

- Package gramEvol in R
- http://cran.r-project.org/web/packages/gramEvol/index.html
Grammar

› Features = Supply temp. setp. \([t-< n>, ..., t +< n>] +\)

\[ \text{Return temp. } [t -< n>, ..., t +< n] + \]

\[ \text{Sys. Power } [t -< n>, ..., t +< n] \]

› Model = SVM(Target ~ Features, \(C = < costs >, \gamma = < gamma >) \)

› \(< n > := 0 | 1 | 2 | 3 | 4 | ... | 20 \)

› \(< costs > := 0.25 | 0.5 | 1 | 2 | 4 | ... \)

› \(< gamma > := 0.25 | 0.5 | 1 | 2 | 4 | ... \)
Grammar

\[\text{Features} = \text{grule}(\text{cbind}(\text{CenteredWindow}(\text{Supply}_\text{temp}_\text{setp}, \text{width} = n),
\text{CenteredWindow}(\text{Return}_\text{temp}, \text{width} = n),
\text{CenteredWindow}(\text{Sys}_\text{Power}, \text{width} = n))))\]

\[\text{Model} = \text{grule}(\text{SVM}(\text{Target} \sim \text{Features}, \ C = \text{costs}, \ \gamma = \text{gamma}))\]

\[n = \text{gvrule}(0:20)\]
\[\text{costs} = \text{grule}(0.25, 0.5, 1, 2, 4)\]
\[\text{gamma} = \text{grule}(0.25, 0.5, 1, 2, 4)\]
Merging the results

› 13 models in total
  - 2 for missing System Supply Temperature data only
    - Change of setpoints behaviour on 1\textsuperscript{st} March
  - One for each other data missing only
  - One for each time-series style forward prediction
  - One for each time-series style backward prediction

› Note:
  - Daylight saving times and time-stamps
Results from all periods combined:
› Competition information and data:

› Competition entry code available:
  - https://github.com/fnoorian/geccoic15/

› Grammatical evolution tutorial in R
  - http://cran.r-project.org/web/packages/gramEvol/vignettes/ge-intro.pdf

› gramEvol package code:
  - https://github.com/fnoorian/gramEvol

› Any questions?
  - Email: farzad.noorian@sydney.edu.au