FLANDERS MAKE
DRIVING INNOVATION IN MANUFACTURING

Pattern-based Time Series Segmentation

Louis Carpentier, Len Feremans, Wannes Meert, Mathias Verbeke

FM@KUL - MPro

Digitized Production enabling end-to-end design-operation

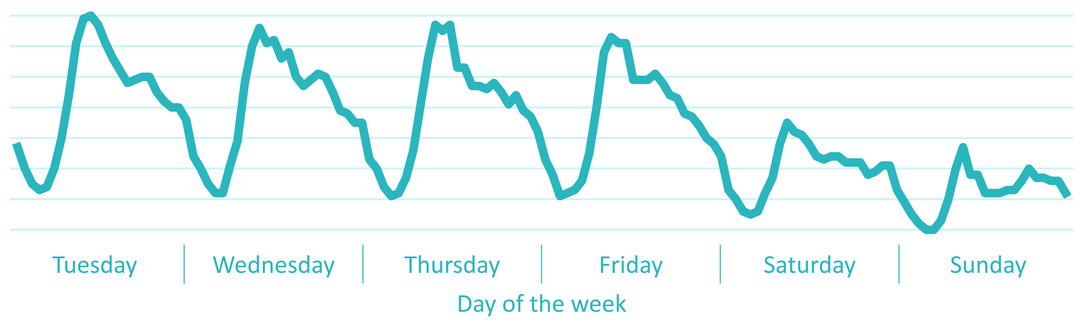
Goal

Partition the time domain in segments which share similar semantics in the application domain, e.g.,
- weekday vs. weekend
- day vs. night
- production of product A vs. product B

⇒ learn about the application domain

Input data

(Multi-variate) time series representing dynamic behavior of a particular real-world phenomenon

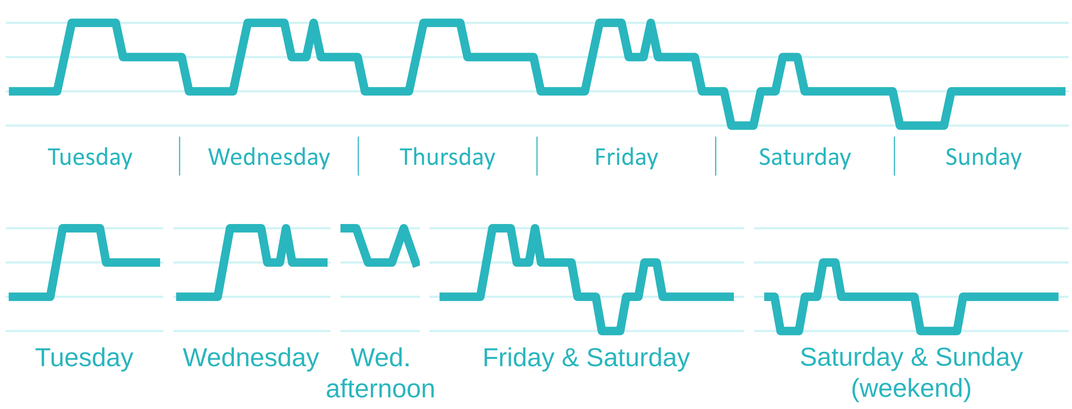


Preprocessing

Discretization using SAX

Segmentation

- Possibly overlapping windows
- Multiple resolutions: window length of a day, two days, an hour, …



Further reading

Frequent pattern mining

Variable-length sequential patterns

- Allow gaps in the window
- Prune with minimal description length
- Mine patterns in every resolution independently

Jaccard similarity

- Filter patterns that cover similar windows

Maximum variance

- Remove patterns that only cover few windows
- Remove patterns that cover many windows (irrelevant for segmentation)

	Tue	Wed	Thu	Fri	Sat	Sun	Support
	X	X	X	X			4 / 6
	X	X (gap)	X	X (gap)			4 / 6
					X	X	2 / 6

	Tue	Wed	Thu	Fri	Sat	Sun
	X	X	X	X		
	X	X (gap)	X	X (gap)		

Covered windows are identical (weekdays)

Pattern-based embedding

$$E(P,W) = \begin{cases} support(P) & pattern\ P\ occurs\ in\ window\ W \\ 0 & otherwise \end{cases}$$

- Support instead of 1 to focus on frequent patterns
- Concatenate embedding matrices of every resolution

	Tue	Wed	Thu	Fri	Sat	Sun
	4/6	4/6	4/6	4/6	0	0
	0	0	0	0	2/6	2/6

Semantic segmentation

Clustering

- Windows with similar embedding vector contain similar patterns
- Cluster columns of embedding matrix = cluster time domain

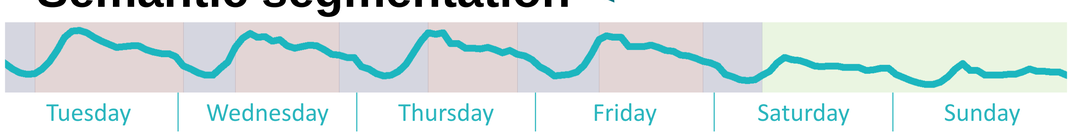
	Tue	Wed	Thu	Fri	Sat	Sun
	4/6	4/6	4/6	4/6	0	0
	0	0	0	0	2/6	2/6

Covered by pattern 1 = weekday

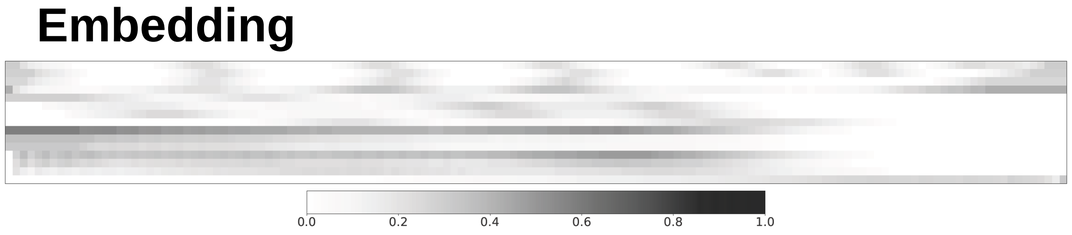
Covered by pattern 2 = weekend

Algorithmic results

Semantic segmentation



Embedding



Key take-aways

- Find intervals in the time series by mining frequent sequential patterns
- Long- and short-term patterns due to multiple resolutions
- Explain semantic segments through the mined patterns

Louis.carpentier@kuleuven.be

For Publication