SCRIMP motif discovery: Matrix Profile methods compute nearest-neighbor distances for subsequences. SCRIMP incrementally approximates the profile and refines it efficiently. Provided via tsmp and wrapped by hmo\_mp().

Objectives: This Rmd demonstrates motif discovery using Matrix Profile with SCRIMP via hmo\_mp("scrimp", ...). SCRIMP is an incremental algorithm for computing the matrix profile. Steps: load packages/data, visualize, set subsequence length and count, fit, detect, evaluate, and plot.

# Install Harbinger (only once, if needed)  
#install.packages("harbinger")

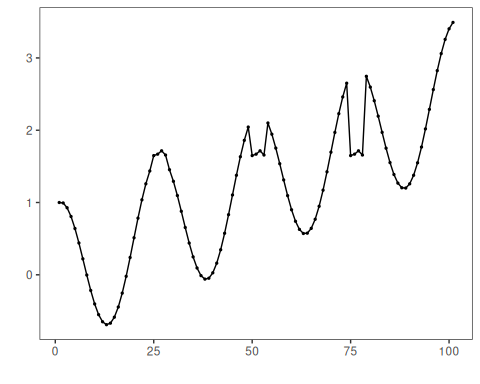
# Load required packages  
library(daltoolbox)  
library(harbinger)

# Load example datasets bundled with harbinger  
data(examples\_motifs)

# Select a simple example time series  
dataset <- examples\_motifs$simple  
head(dataset)

## serie event  
## 1 1.0000000 FALSE  
## 2 0.9939124 FALSE  
## 3 0.9275826 FALSE  
## 4 0.8066889 FALSE  
## 5 0.6403023 FALSE  
## 6 0.4403224 FALSE

# Plot the time series  
har\_plot(harbinger(), dataset$serie)



# Define Matrix Profile (SCRIMP) motif model  
# - second arg: subsequence length (window)  
# - third arg: number of motifs to retrieve  
 model <- hmo\_mp("scrimp", 4, 3)

# Fit the model  
 model <- fit(model, dataset$serie)

# Detect motifs  
 detection <- detect(model, dataset$serie)

## Finished in 0.07 secs

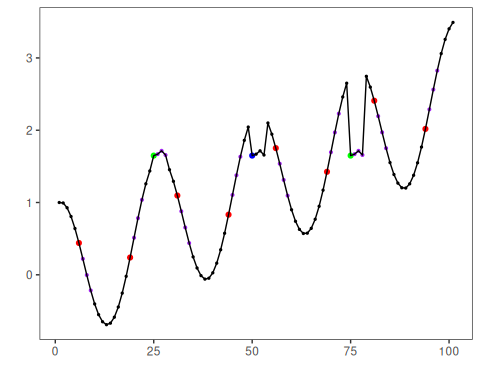
# Show only timestamps flagged as events  
 print(detection |> dplyr::filter(event==TRUE))

## idx event type seq seqlen  
## 1 6 TRUE motif 3 4  
## 2 19 TRUE motif 2 4  
## 3 25 TRUE motif 1 4  
## 4 31 TRUE motif 3 4  
## 5 44 TRUE motif 2 4  
## 6 56 TRUE motif 3 4  
## 7 69 TRUE motif 2 4  
## 8 75 TRUE motif 1 4  
## 9 81 TRUE motif 3 4  
## 10 94 TRUE motif 2 4

# Evaluate detections against ground-truth labels  
 evaluation <- evaluate(model, detection$event, dataset$event)  
 print(evaluation$confMatrix)

## event   
## detection TRUE FALSE  
## TRUE 2 8   
## FALSE 1 90

# Plot detections over the series  
 har\_plot(model, dataset$serie, detection, dataset$event)



References - Yeh, C.-C. M., et al. (2016). Matrix Profile I/II: All-pairs similarity joins and scalable time series motif/discord discovery. IEEE ICDM. - Tavenard, R., et al. (2020). tsmp: The Matrix Profile in R. The R Journal. <doi:10.32614/RJ-2020-021>