IEEE International Conference on Bioinformatics & Bioengineering
2015

False Alarm Suppression in Early Prediction of Cardiac Arrhythmia

Shoumik Roychoudhury, Mohamed F. Ghalwash, Zoran Obradovic

Presented by: Zoran Obradovic

zoran.obradovic@temple.edu

Center for Data Analytics and Biomedical Informatics

TEMPLE UNIVERSITY

Introduction

- Alarm fatigue among care providers inside ICUs due to the high percentages of bedside monitor false alarms, has been identified as one of the top 10 medical hazards.

- Alarm fatigue results in desensitization among care providers, which ultimately can lead to lower standards of care to patients and also result in fatal consequences.

Introduction

- Goals
  - Suppress high false alarm rates from bedside monitors in intensive care units.
  - Improve the response time of the medical personnel in the event of life-threatening arrhythmia alarms by providing early prediction of alarms.
  - Reduce alarm fatigue among care providers by early false alarm suppression.

- Objective
  - Develop a cost-sensitive prediction model to suppress false alarms while keeping true alarm detection rates high.
  - Build a model for early prediction of alarm events.

Contribution

	Li et al.	Behar et al.	EDSC	Our approach
Intrepretabilty	×	×	✓	✓
Earliness	×	×	✓	✓
Uncertainty	×	×	×	✓
False alarm suppression	✓	✓	✓	✓
Cost-sensitive	×	×	×	✓

- Provide more **accurate** prediction (**high true alarm detection and false alarm suppression**) than the state-of-the-art methods on arrhythmia alarms.

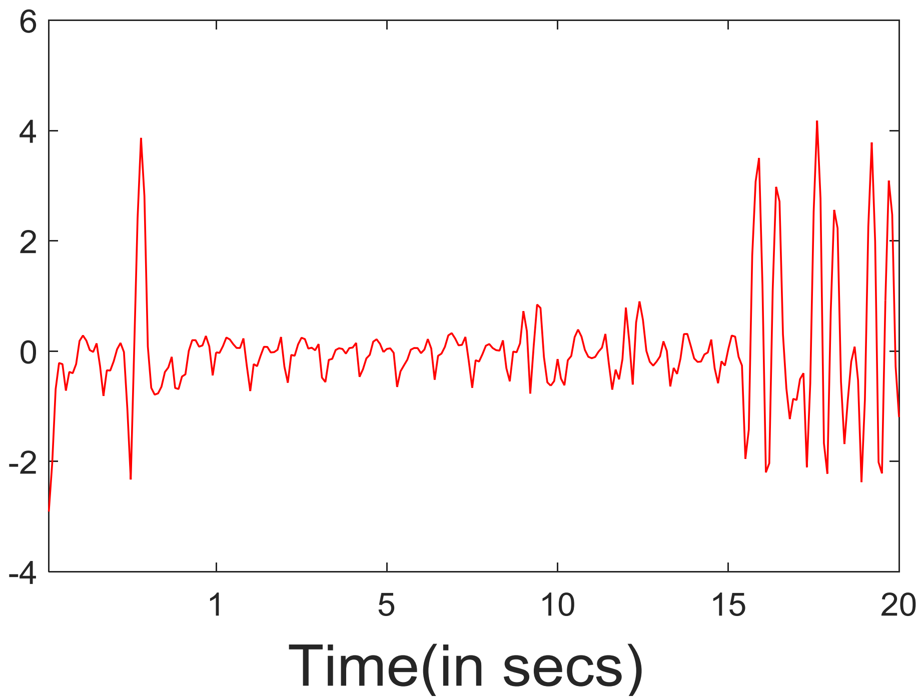
- Provide **interpretable** results in order to explain the rationale of the prediction, whereas all other published methods are black-box.

- Provide **early** prediction before the alarm happens, which helps practitioners respond earlier.

- Provide a **cost-sensitive** model to achieve the desired level of false alarm suppression rate.

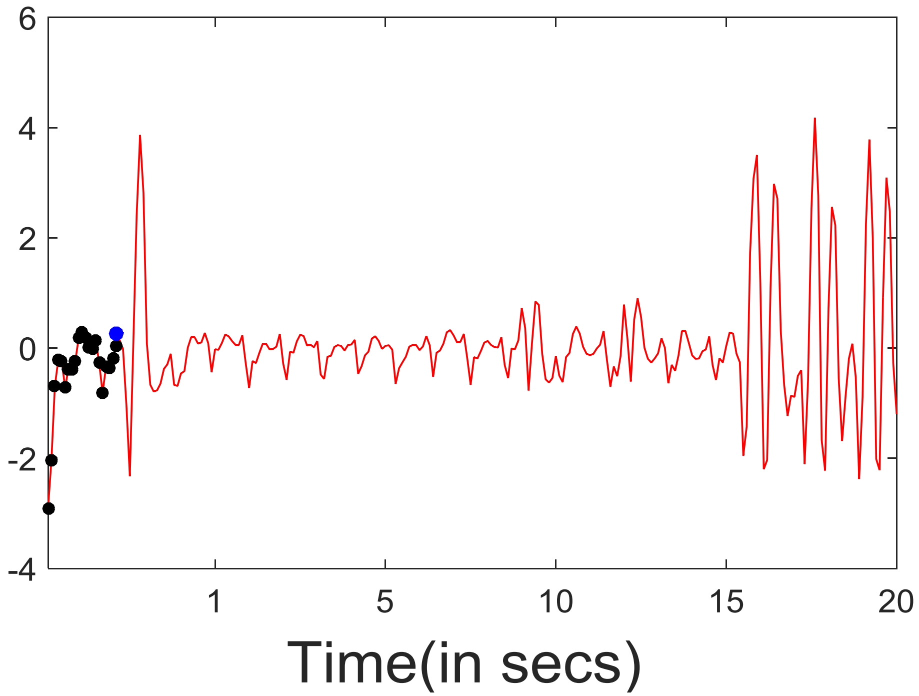
Early prediction of time series

- The main objective is to predict the label of the time series as the signal is progressively recorded and before the alarm event occurs.
- Example:
  - Red signal represents a True alarm signal.
  - 20-second analysis window prior to the alarm is considered. The alarm occurs at the 21^{st} second. The observed window below represents 20 seconds prior to the alarm.



EDSC (Xing et al. 2011)

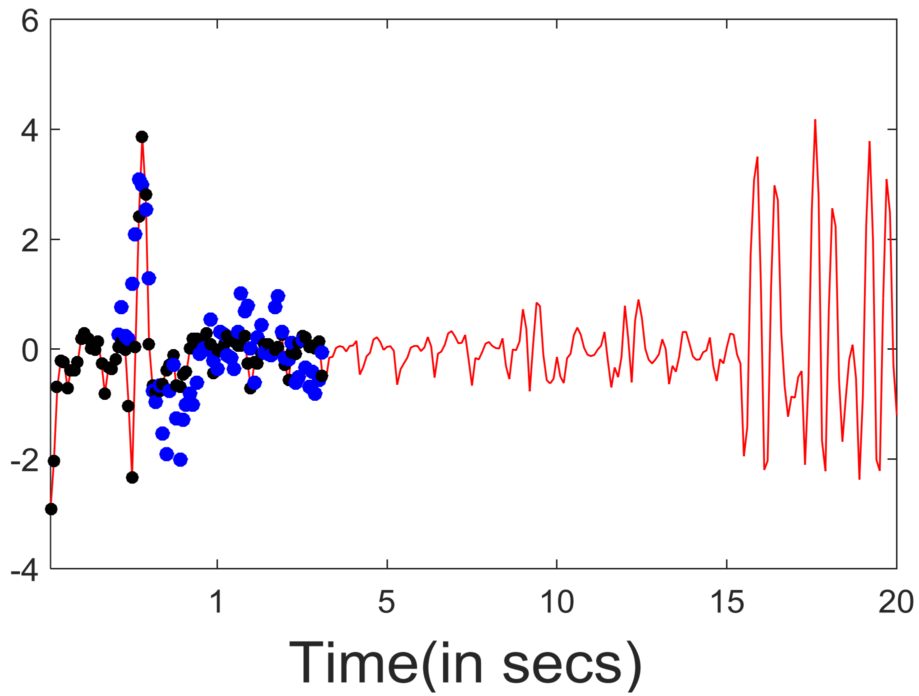
- Interpretable early classification model (EDSC) extracts local discriminative patterns (shapelets) from the time series in order to characterize the target class.



- Black marks indicate the algorithm is reading an unknown time series.

EDSC

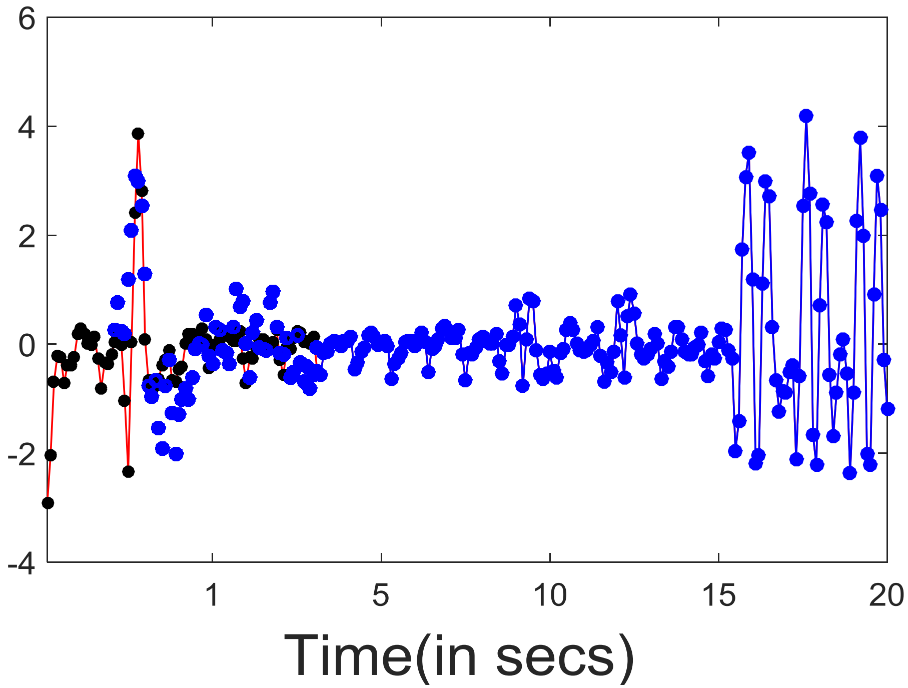
- Interpretable early classification model (EDSC) extracts local discriminative patterns (shapelets) from the time series to characterize the target class.



- Blue shapelet S is a good match at time step 4. It is extracted during the learning phase and is identified as a characteristic pattern for a false alarm event.

EDSC

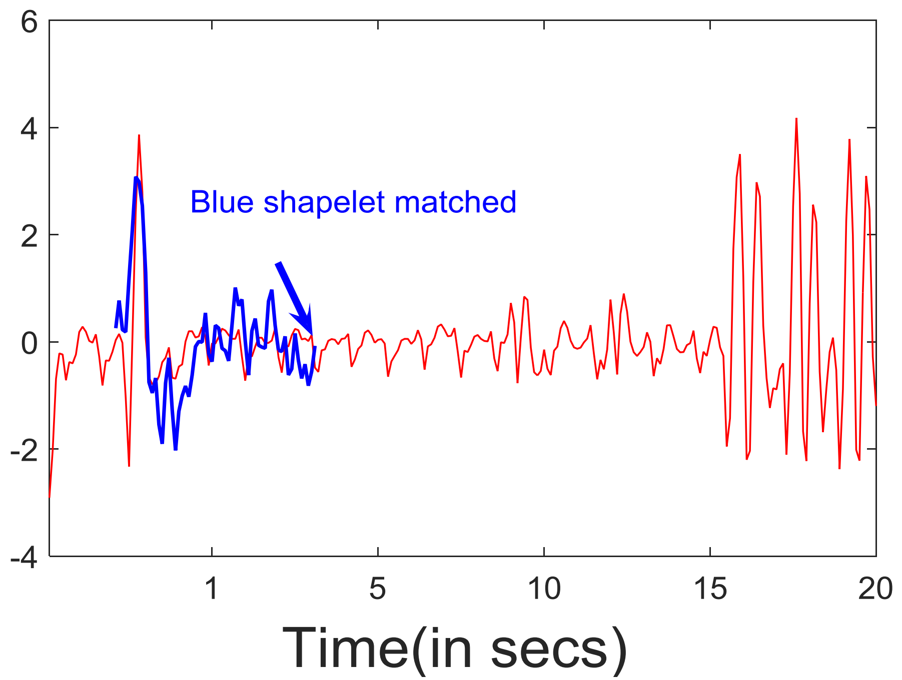
- Interpretable early classification model (EDSC) extracts local discriminative patterns (shapelets) from the time series in order to characterize the target class.



- The red signal is misclassified as a blue class at time step 4.

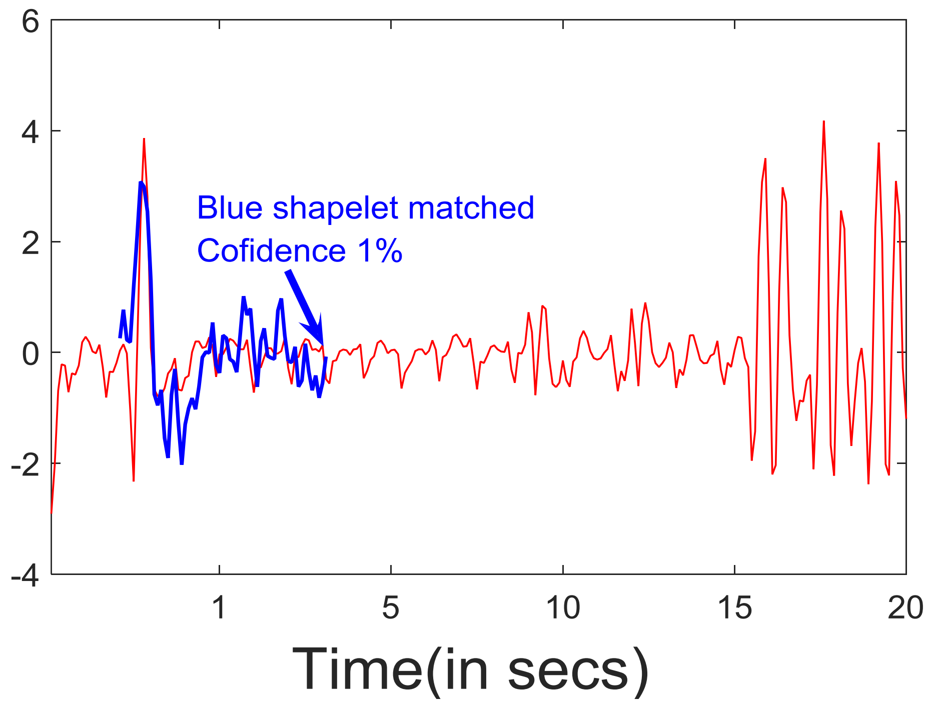
Drawbacks of EDSC

- Signals from different classes may be similar to each other in early phases.
- Extracted shapelets might have similar pattern which can lead to wrong predictions.
- e.g. a true alarm signal might match a false alarm shapelet.
- EDSC method lacks uncertainty (confidence) estimation of predicted labels.



MEDSC-U (Ghalwash et al. 2013)

- MUDSC-U extends EDSC to produce interpretable early classification of time series with uncertainty estimates.
- The uncertainty for the matched subsequence (distance measurement) is computed.

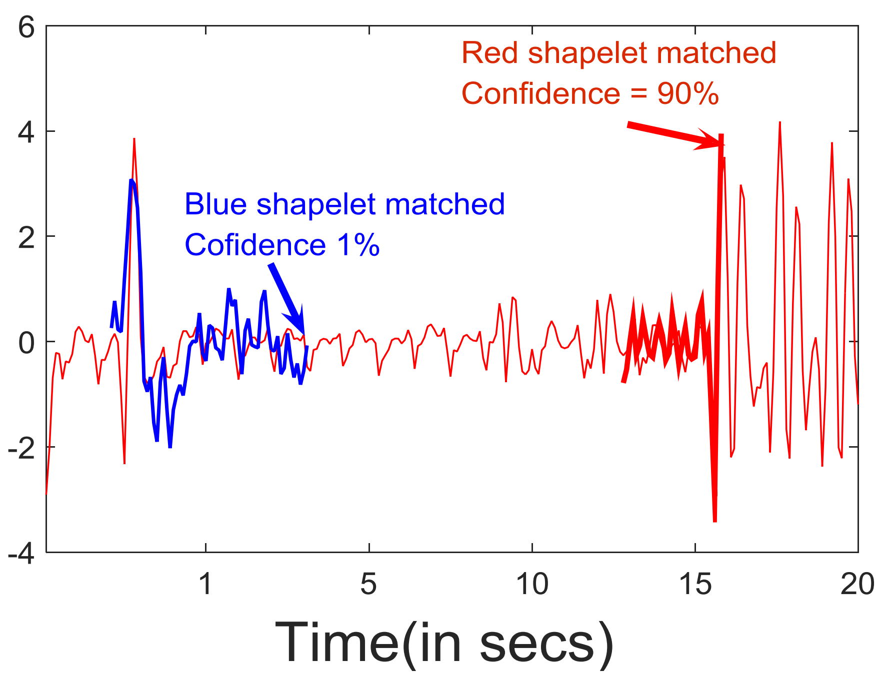


- MEDSC-U computes low confidence for a blue shapelet match.

MEDSC-U

- For a ECG signal *T* with unknown label (true or false alarm), the Z-normalized Euclidean distance between the observed signal and all extracted shapelets (from training phase) is computed.
- The distance between T and S contains uncertainty which is defined as a random variable *d = dist(s,T) +ε*

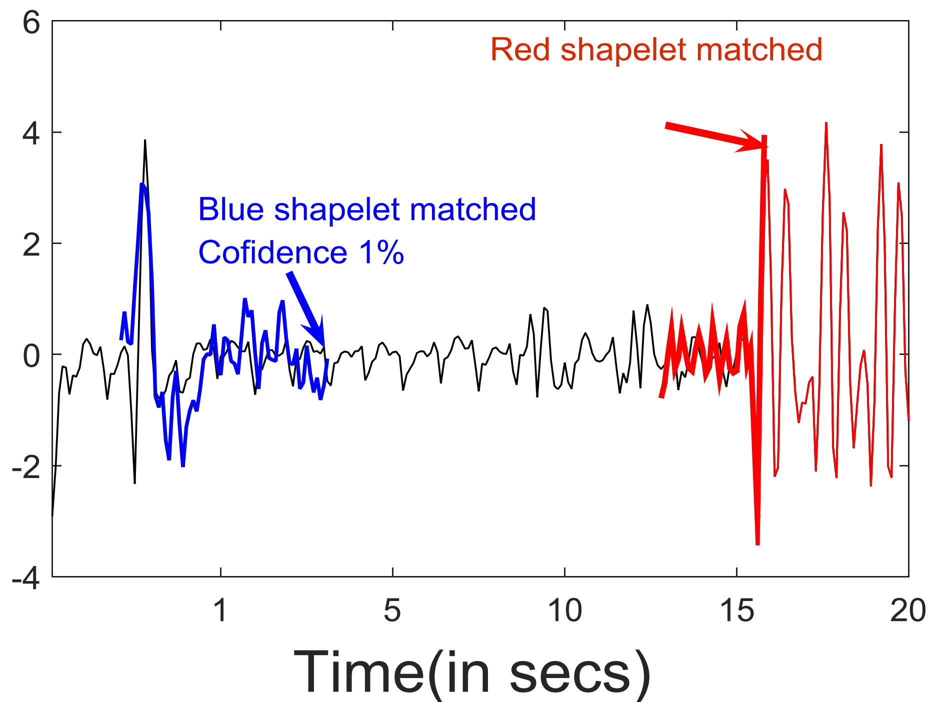
Confidence estimate : $C_S^c \geq \frac{(\delta - dist(s,T))^2}{\sigma^2 + (\delta - dist(s,T))^2} \times Precision(S)$



1-MEDSCU

Cost-sensitive early prediction using confidence estimates

- **The hybrid model** between **MEDSC-U** and **EDSC** methods, utilizes **high confidence** level for predicting **false alarm** and predicts a **true alarm** as soon as a **match** is found.
- The computed **confidence** C_S^c is compared to a predefined class specific confidence threshold value.



Learning Phase

Extract subsequences (shapelets) that provides classification of unknown time series as early as possible.

- **Shapelet extraction**: Extract all shapelets of different lengths, where for each shapelet a distance threshold is learned such that the shapelet discriminates between classes.
- **Ranking shapelets**: Assigns a score to each shapelet that incorporates both the earliness and the accuracy. The earliness defines how early, on average, the shapelet matches the target time series.
- **Pruning Phase**: Prunes the shapelets by selecting the top shapelets that cover the entire dataset.

A window of length l is slided over the time series T to extract all subsequences $\{h_1, h_2, \ldots h_{L-l+1}\}$ of length l.

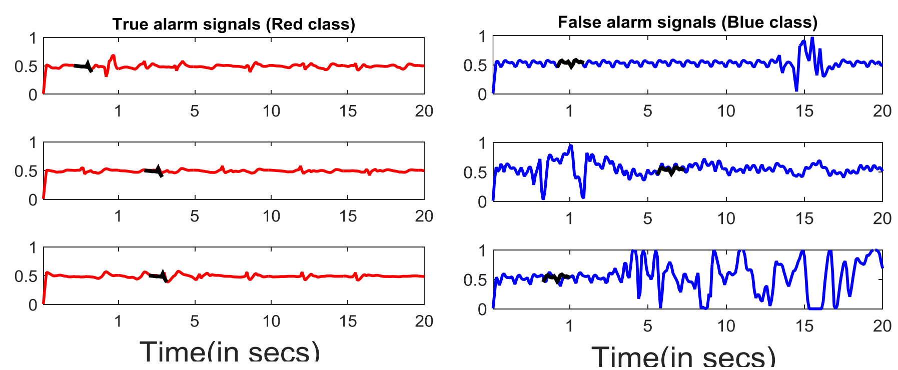
Shapelet $S = (s, l, \delta, c)$ where,

s : time series subsequence,

l : length of subsequence,

c : the alarm label of the subsequence

δ : distance threshold which needs to be learned.



Shapelets (in black) from true alarm (red) and false alarm (blue) Classes.

Experiments: Data and Setup

- ECG lead II data
  - Two types of critical arrhythmia alarm datasets from MIMIC II version 3 repository.
    - Asystole (ASYS)
    - Ventricular tachycardia (VTACH)

Dataset	Total alarms	True alarms(%)	False alarms(%)
ASYS	261	40(15.3%)	221(84.7%)
VTACH	629	227(36.09%)	402(63.91%)

- Setup
  - True alarms are randomly partitioned into 4 disjoint groups of N/4 examples.
  - N/4 false alarms are randomly selected for each of 4 groups.
  - All methods (the proposed and the baseline methods) are trained using 2 groups (N/2 true and N/2 false alarms) and tested on the remaining 2 examples.
  - This is repeated 20 times

Experiments: Measures and Baselines

- Evaluation measures
  - True alarm detection rate (TAD) = sensitivity
  - False alarm suppression rate (FAS) = specificity
  - Earliness
  - Weighted balanced accuracy (WAcc)
    - Misclassification do not have equal weights. False negative (FN) are more costly than false positives(FP).

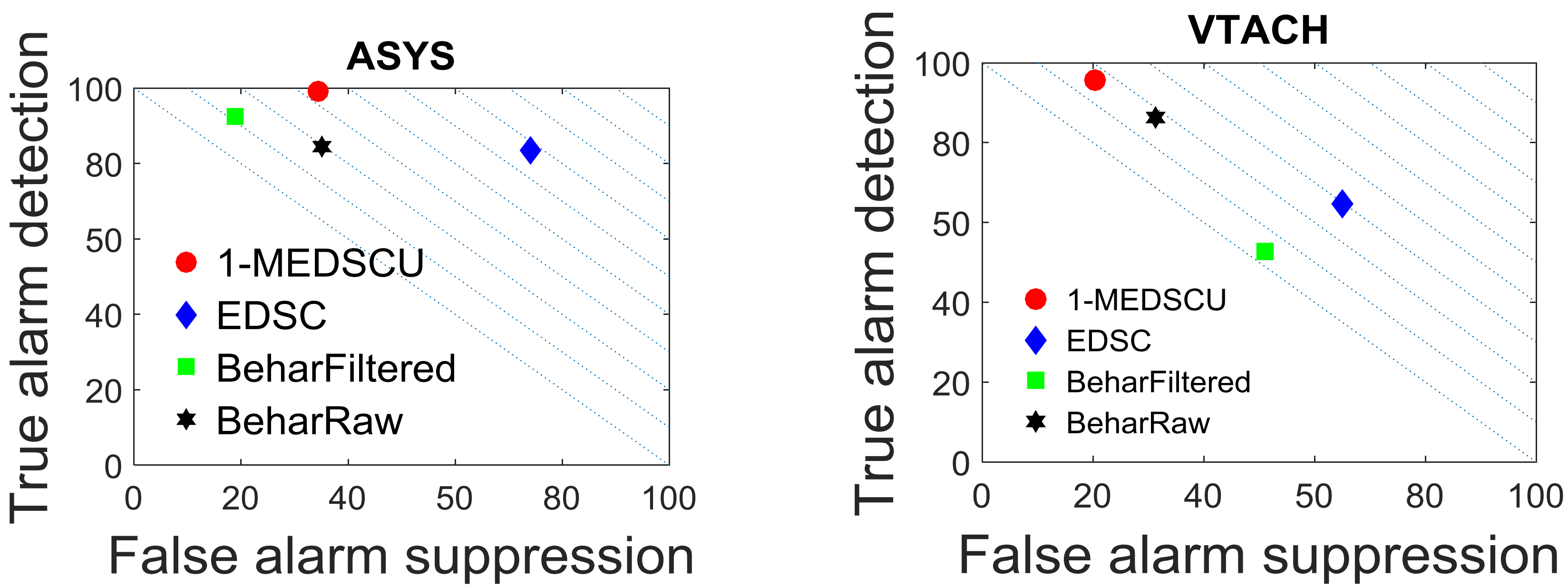
$$WAcc = \frac{WSens + Spec}{2} \, where,$$

$$WSens = \frac{TP}{TP + (1+\beta^2)FN} \, where \; \beta = \{2,3\} \;, \qquad Spec = \frac{TN}{TN + FP}$$

- Baseline methods
  - Behar et al. 2013, Black box method using feature extraction and SVM .
  - EDSC (Xing et al. 2011), Early classification using Euclidean distance measure.

TAD Vs. FAS

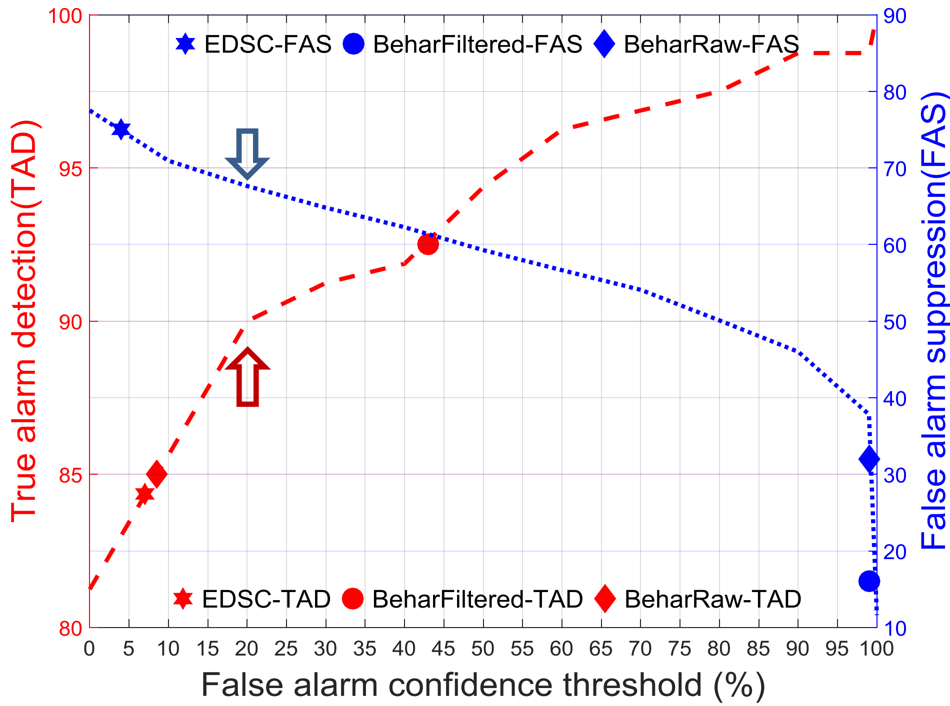
- Objective: Achieve high FAS (X-axis) while keeping near 100% TAD (Y-axis).
- Increasing value in X-axis indicates high false alarm suppression and increasing value in Y-axis indicates high true alarm detection.
- The markings indicate performance of each model for both TAD and FAS



- The upper right hand corner in the figure is the ideal result (100 % FAS and 100 % TAD).
- The proposed method (red circle) outperforms all baseline methods in terms of TAD in both datasets.
- Baseline methods are better in terms of FAS however they make lot of false negative errors.

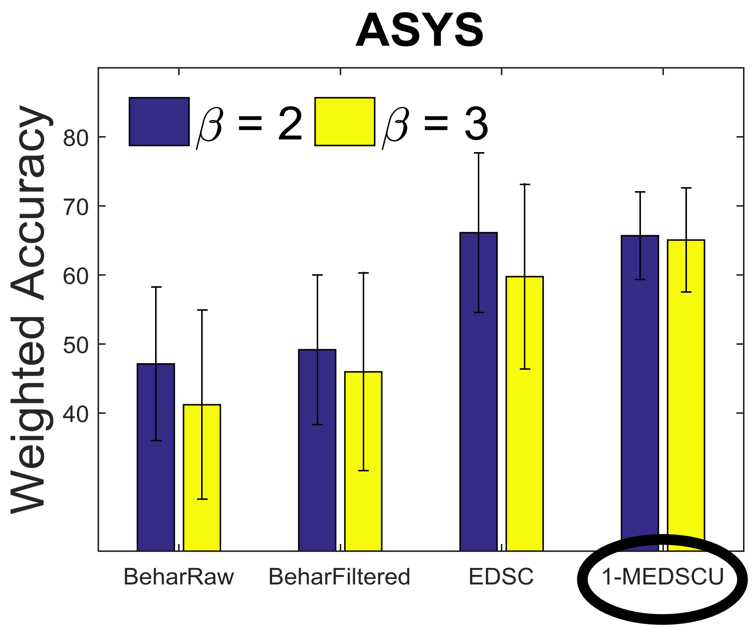
Controlling False Alarm Suppression Rate

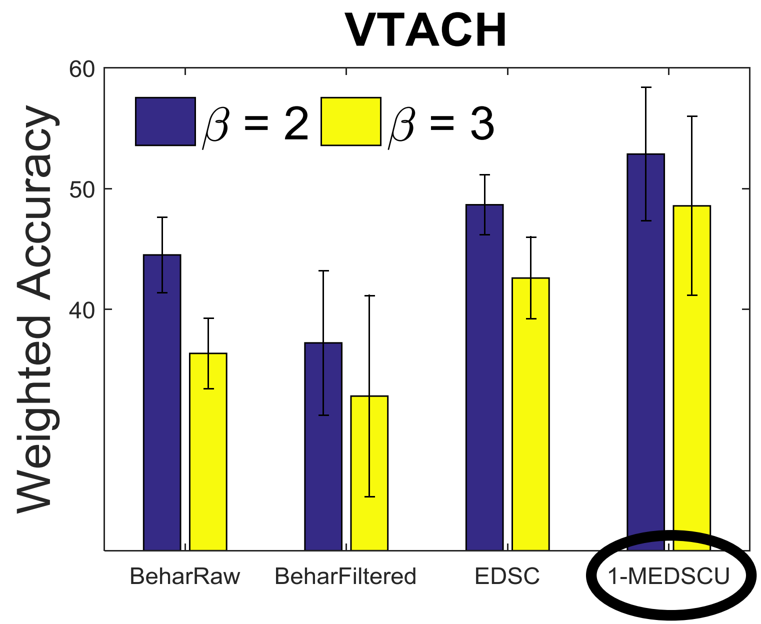
- Baseline methods are **rigid** in terms of true alarm detection and false alarm suppression (red and blue marks).
- 1-MEDSCU has a **configurable** framework where users can set-up a desired TAD rate (dashed red line) by controlling the false alarm confidence threshold (x-axis). The blue dotted line is the varying false alarm suppression.
- e.g 1-MEDSCU can achieve a 68% false alarm suppression (blue arrow) by dropping the true alarm detection rate to 90% (red arrow) in case of ASYS alarms.



Varying false alarm confidence threshold for ASYS

Weighted balanced accuracy (WAcc)





- False positive and false negative errors should not have equal weights. False negative errors are more costly than false positive errors. False negatives can cause patient death (missed true alarm).

$$WAcc = \frac{WSens + Spec}{2} \, where, \qquad WSens = \frac{TP}{TP + (1+\beta^2)FN} \, where \; \beta = \{2,3\}$$

$$Spec = \frac{TN}{TN+FP}$$

- False negative errors are penalized more using the WAcc evaluation measure. For example $\beta = 2, 3$ penalizes false negative error more than $\beta = 1$, which represents traditional balanced accuracy.
- 1-MEDSCU outperforms all other methods in terms of WAcc.

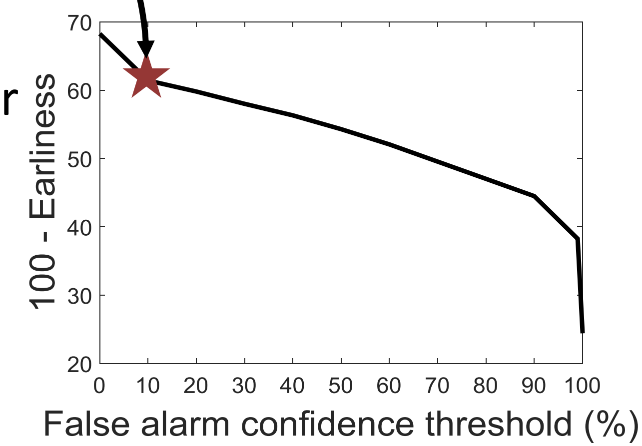
Earliness

- 1-MEDSCU provides earlier predictions than Behar et al., but EDSC predicts even earlier.
- We report 100-earliness (Larger the better).

Dataset	Behar(Raw)	Behar(Filtered)	EDSC	1-MEDSCU
ASYS	0	0	62.8±6.27	38.39±9.05
VTACH	0	0	59.9±11.71	39.96±9.34

- However, 1-MEDSCU is a **configurable** model, we can improve the earliness by varying the false alarm confidence threshold. End users have the option of choosing their desired level of earliness.
- Figure shows how earliness is improved for lower values of false alarm confidence threshold.

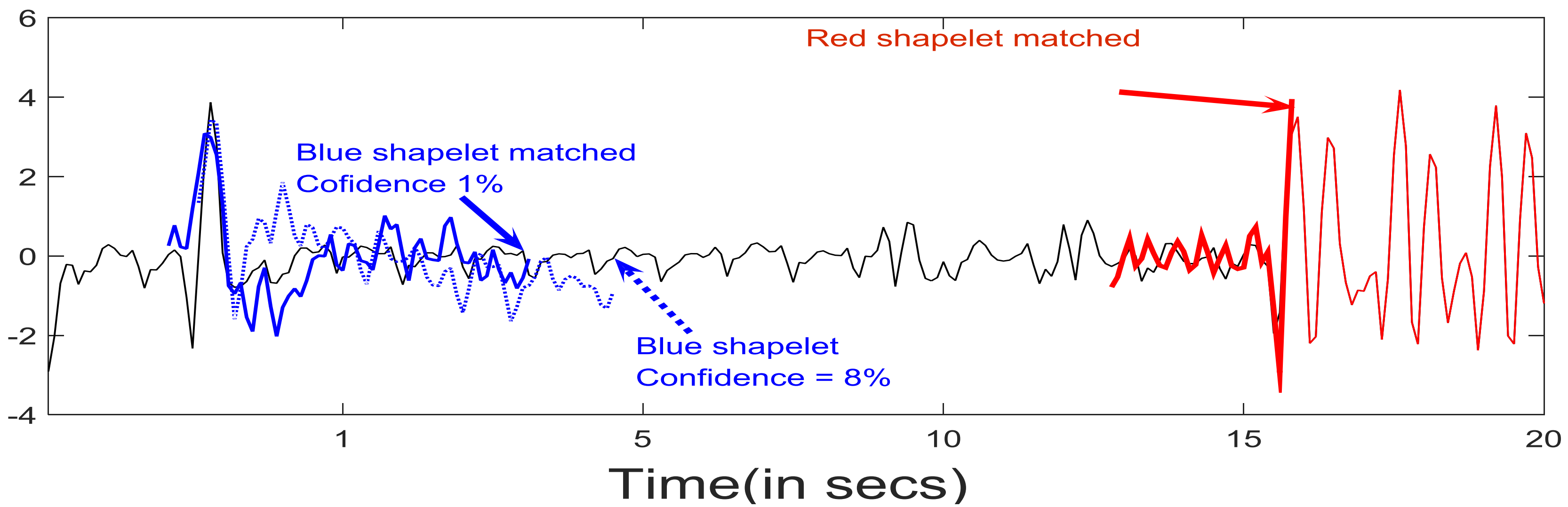
1-MEDSCU can predict earlier than EDSC when false alarm confidence threshold is smaller than 20%.



100- Earliness values for varying false alarm confidence threshold for ASYS

Interpretability

- 1-MEDSCU is an **interpretable** method, thus domain experts can **visually verify** the actual reason for false or true alarms.
- The exact time of the pattern can also be identified.
- **Domain experts prefer** such methods compared to **black-box classification methods**.



The true alarm is wrongly classified as a false alarm by EDSC at time 4.

It is correctly classified as a true alarm by 1-MEDSCU at time 16.

Summary

- We were able to achieve a moderate percentage of FAS while keeping high rate of early TAD predictions.
- The proposed approach has outperformed the state-of-the-art methods in terms of weighted accuracy.
- We can achieve higher FAS rates by sacrificing TAD rate.
- In addition, the proposed method provides not only accurate results but also interpretable and early.

- Future work
  - Currently, 1-MEDSCU works for univariate time series only. We will extend it to multivariate time series in order to improve the performance of the model.
  - We will also investigate the problem of false alarm suppression in other medical domains.

Thank you