TimeSeg: An Information-Theoretic Segment-Wise Explainer for Time-Series Predictions

Hwijin Kim, Jaeho Kim, Changhee Lee

1. Motivation

Challenges in Time-Series Explanation.

- Time-series models exhibit *dynamically evolving patterns* and *complex temporal dependencies*
- Many real-world models are *strict black-boxes*: no access to *internal states*
- Generated explanations must be *interpretable* while *preserving predictive information*

Limitations of Prior Approaches.

Point-wise Methods (e.g., TimeX++, WinIT):

- 👍 Precise localization
- 👎 Fragmented explanations
- 👎 Lack interpretability

Patch-wise Methods (e.g., LIMESegment):

- 👍 Contiguous segments
- 👎 Fixed-length limitations
- 👎 Misaligned boundaries

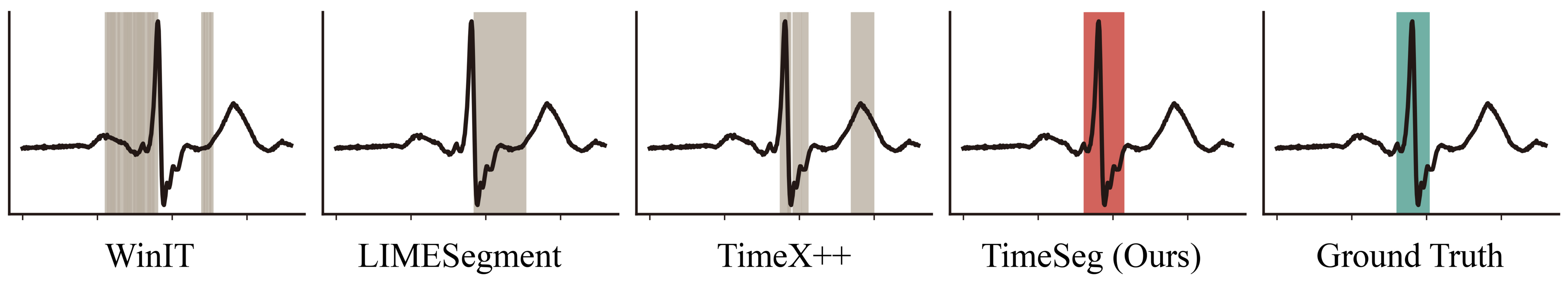


Figure 1: Comparison of generated explanations for MIT-ECG dataset.

Key Insight.

⇒ Need *variable-length segment-wise explanations* that are both *temporally coherent* and *precisely localized*

2. Problem Formulation

Notation and Setup.

- Black-box classifier: $g_\theta : \mathbb{R}^T \to \mathbb{R}^C$ maps univariate time-series $\mathbf{x} = (x_1, \dots, x_T)$ to C classes
- Segment-index variable: $s := (t^s, t^e)$ where $t^s \le t^e$ denote start and end time points
- Goal: Find K non-overlapping segments $\mathbf{s}_{1:K} = (s_1, \dots, s_K)$ most predictive of target

Optimal Segment-wise Explainer.

- The optimal segment-wise explainer $\mathcal{E}^*$ is defined as:

$$\mathcal{E}^* = \arg\max_{\mathcal{E}} I\left(g_\theta(\mathbf{X}); \mathbf{X}_{\mathbf{s}_{1:K}}\right) - \lambda J(\mathbf{s}_{1:K}) \quad \text{subject to} \quad \mathbf{s}_{1:K} \sim \mathcal{E}(\mathbf{X})$$

- $I(\cdot;\cdot)$: Mutual information measuring predictive power of selected segments
- $J(\cdot)$: Regularizer for segmentation complexity (e.g., segment length)
- $\lambda \in \mathbb{R}^+$: Trade-off coefficient between informativity and simplicity

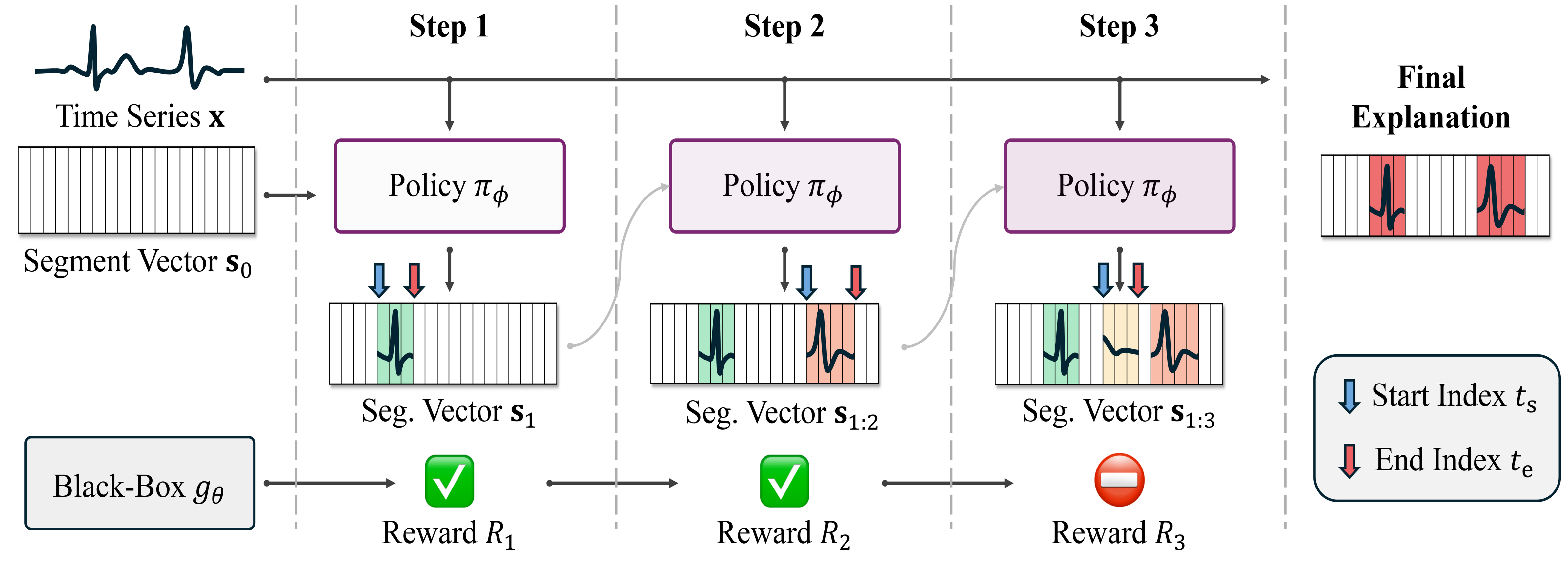
Key Challenges.

- Estimating joint MI between arbitrary segment collections and target is *intractable*
- Combinatorial search over all possible segments grows *exponentially* $\mathcal{O}(2^T)$

Takeaway.

⇒ Decompose joint MI into *conditional MI (CMI)* via chain rule, formulate as *sequential decision process*

3. Method: TimeSeg



Sequential MI Decomposition.

- Decompose joint MI into sum of CMI terms:

$$I\left(g_\theta(\mathbf{X}); \mathbf{X}_{\mathbf{s}_{1:K}}\right) = \sum_{k=1}^{K} I\left(g_\theta(\mathbf{X}); \mathbf{X}_{\mathbf{s}_k} | \mathbf{X}_{\mathbf{s}_{1:k-1}}\right)$$

RL-based Segment Selection.

- **At each step**: Given state $(\mathbf{x}, \mathbf{s}_{1:k-1})$, policy π_ϕ takes action to select next segment $s_k = (t^s, t^e)$
- **Reward**: CMI approx with sparsity penalty:

$$R_k := \underbrace{\mathbb{E}_{p_\theta(y|\mathbf{x})}\left[\log p_\theta(y|\mathbf{x}_{\mathbf{s}_{1:k}}) - \log p_\theta(y|\mathbf{x}_{\mathbf{s}_{1:k-1}})\right]}_{\text{cross-entropy gap (CMI, } r_\theta)} - \lambda \underbrace{\tfrac{1}{T}\|\mathbf{m}_k\|_1}_{\text{sparsity penalty}}$$

Two-step Policy Factorization.

- Factorize into *start* and *end* policy ($t^s \le t^e$):

$$\pi_\phi(\mathbf{s}|\cdot) = \underbrace{\pi_{\phi_s}(t^s|\cdot)}_{\text{where-to-start}} \cdot \underbrace{\pi_{\phi_e}(t^e|t^s,\cdot)}_{\text{where-to-end}}$$

- Policy π_{ϕ^e} *masks invalid indices* before softmax
- Optimize via *PPO* with actor-critic framework

Adaptive Termination.

- Terminate when segment s_k provides negligible CMI gain, below threshold τ:

$$\frac{r_\theta(\mathbf{x}_{\mathbf{s}_k}, \mathbf{x}_{\mathbf{s}_{1:k-1}})}{\mathrm{CE}_{p_\theta}[y|\mathbf{x}, y|\mathbf{x}_{\mathbf{s}_{1:k-1}}]} \le \tau \quad \Rightarrow \quad \textbf{Stop}$$

- Enables *instance-specific* number of segments K

4. Experiments

Datasets.

Synthetic: SeqComb, FreqShapes, LowVarDetect **Real-world**: MIT-ECG, Epilepsy, Wafer, GunPoint

Results.

Method	F1 ↑	IoU ↑	Cont. ↓
IG*	0.589±0.006	0.435±0.007	0.056±0.007
Dynamask	0.353±0.023	0.221±0.018	0.072±0.016
WinIT	0.241±0.062	0.147±0.042	0.133±0.014
LIMESegment	0.491±0.068	0.359±0.063	0.006±0.001
TimeX++*	0.593±0.146	0.460±0.142	0.016±0.002
TimeSeg	**0.739±0.016**	**0.621±0.021**	**0.006±0.000**

Table 1: MIT-ECG (with ground-truth).

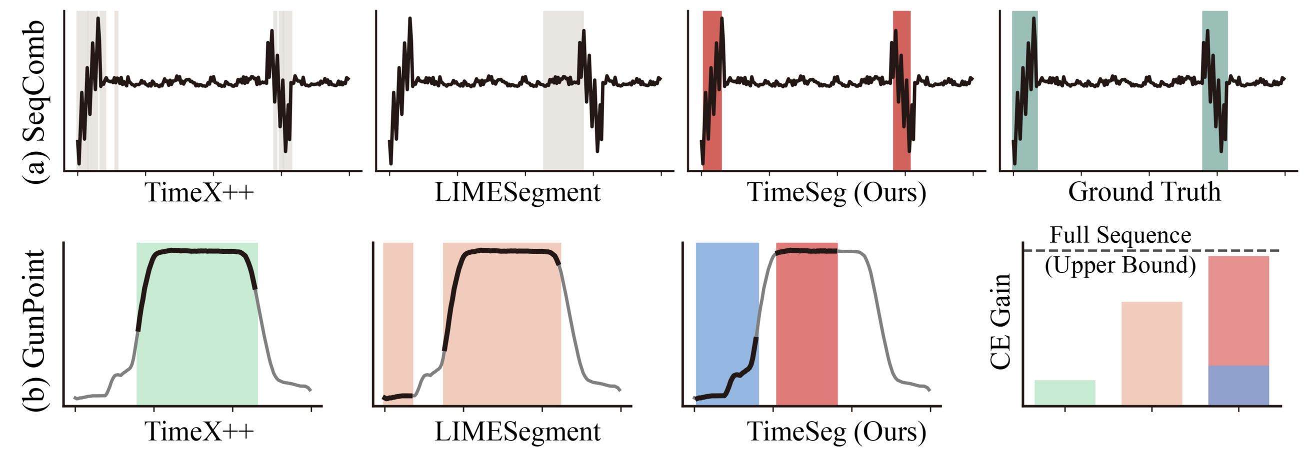


Figure 2: Case Study (a) with, (b) without ground truth.

Key Findings.

- **Avg. rank**: F1/IoU *1.5* (4 datasets), Fidelity *1.5* (4 datasets), Contiguity *1.57* (7 datasets)
- **Case study**: TimeSeg provides segments that are *accurately localized*, and *maximally informative*