STRADA: Training/Benchmark System Configuration and Model Parameters

CPU	AMD EPYC 7302
GPU	NVIDIA RTX A5000
RAM	128 GB
OS	Ubuntu 18.04.5 LTS
CUDA Version	11.6

TABLE I: Training and benchmark system configuration used in *STRADA: A Synergising Transformer and DBSCAN-Based Approach to Time-Series Anomaly Detection in Dynamically-Configured Systems*.

Dataset	l_{min}	l_{max}
machine-1-1	4	128
HLT Datasets	4	256

(a) MERLIN

Parameter	Value
ϵ	3
min. sample size	4

(b) T-DBSCAN

Parameter	Model	Value
Sequence Length	All	10
Encoder Layers	All	1
Feed-Forward Layers	All	2
Hidden Units	All	64
Batch Size	All	128
Training Epochs	All	5
Learning Rate	All	10^{-4}
SPOT q	machine-1-1	
	HLT Datasets	10^{-2}

(c) TranAD/STRADA-TranAD

TABLE II: Overview of model parameters used in *STRADA: A Synergising Transformer and DBSCAN-Based Approach to Time-Series Anomaly Detection in Dynamically-Configured Systems.*

Parameter	Model	Value
Sequence Length	machine-1-1	64
	HLT Datasets	16
Encoder Layers	machine-1-1	3
	HLT Datasets	1
Decoder Layers	machine-1-1	2
	HLT Datasets	4
Attention Heads	machine-1-1	8
	HLT Datasets	4
d_{model}	machine-1-1	512
	HLT Datasets	576
d_{ff}	machine-1-1	2048
	HLT Datasets	2944
Learning Rate	machine-1-1	10^{-4}
	HLT Datasets	9.7×10^{-6}
Training Epochs	All	4
SPOT q	machine-1-1	10^{-8}
	HLT Datasets	2.5×10^{-3}

(d) Informer-MSE/STRADA-MSE

Parameter	Model	Value
Sequence Length	All	64
Encoder Layers	All	3
Decoder Layers	All	2
Attention Heads	All	8
d_{model}	All	512
d_{ff}	All	2048
μ	machine-1-1	0.12
	HLT Datasets	0.05
Batch Size	All	64
Learning Rate	All	10^{-4}
Training Epochs	All	4
SPOT q	machine-1-1	10^{-8}
	HLT Datasets	8×10^{-3}

(e) Informer-SMSE/STRADA-SMSE

TABLE II: Overview of model parameters used in *STRADA: A Synergising Transformer and DBSCAN-Based Approach to Time-Series Anomaly Detection in Dynamically-Configured Systems* (continued).