ARES: Anomaly Recognition Model For Edge Streams Online Appendix

Model	DARPA		UNSW-NB15		ISCX2012		CIC-IDS2017		CTU-13					
									Scenario 1		Scenario 10		Scenario 13	
	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP
MIDAS-F	0.961 ± 0.005	0.979 ± 0.003	0.793 ± 0.003	0.337 ± 0.004	0.974 ± 0.004	0.287 ± 0.031	**0.997 ± 0.000**	**0.969 ± 0.014**	0.835 ± 0.037	0.077 ± 0.014	**0.924 ± 0.020**	0.530 ± 0.078	**0.934 ± 0.050**	0.135 ± 0.052
AnoEdge-L	0.910 ± 0.007	0.955 ± 0.003	0.750 ± 0.021	0.443 ± 0.033	0.869 ± 0.038	0.097 ± 0.027	0.954 ± 0.006	0.486 ± 0.028	0.479 ± 0.007	0.018 ± 0.001	0.692 ± 0.006	0.181 ± 0.010	0.293 ± 0.009	0.010 ± 0.000
SLADE-H	0.900 ± 0.093	**0.912 ± 0.091**	**0.911 ± 0.118**	**0.757 ± 0.191**	**0.999 ± 0.000**	**0.985 ± 0.022**	**0.731 ± 0.226**	0.215 ± 0.178	0.564 ± 0.077	0.018 ± 0.005	0.399 ± 0.159	0.099 ± 0.020	0.564 ± 0.077	0.018 ± 0.005
ARES-Static	**0.991 ± 0.005**	**0.994 ± 0.003**	**0.984 ± 0.003**	**0.885 ± 0.016**	**0.999 ± 0.000**	0.945 ± 0.010	0.984 ± 0.001	0.811 ± 0.017	**0.970 ± 0.003**	**0.296 ± 0.030**	**0.945 ± 0.024**	**0.701 ± 0.088**	**0.928 ± 0.122**	**0.282 ± 0.119**
ARES-Dynamic	**0.989 ± 0.007**	**0.994 ± 0.003**	0.877 ± 0.009	0.500 ± 0.030	**0.999 ± 0.000**	**0.960 ± 0.018**	0.984 ± 0.001	0.816 ± 0.013	**0.947 ± 0.027**	**0.209 ± 0.073**	**0.924 ± 0.022**	**0.673 ± 0.091**	**0.851 ± 0.042**	**0.098 ± 0.040**

Table 1: Comparison of ARES **against competing methods by evaluating performance on subsampled versions of the test set for each dataset.**

Model	DARPA		UNSW-NB15		ISCX2012		CIC-IDS2017		CTU-13					
									Scenario 1		Scenario 10		Scenario 13	
	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP	ROC-AUC	AP
GNN-RecErr	0.304 ± 0.000	0.034 ± 0.000	0.629 ± 0.000	0.247 ± 0.000	0.559 ± 0.000	0.059 ± 0.000	0.419 ± 0.000	0.126 ± 0.000	0.394 ± 0.000	0.022 ± 0.000	0.222 ± 0.000	0.041 ± 0.000	0.487 ± 0.000	0.023 ± 0.000
GCN+HST	0.637 ± 0.101	0.806 ± 0.039	**0.986 ± 0.001**	**0.894 ± 0.004**	0.969 ± 0.009	0.347 ± 0.085	0.881 ± 0.061	0.352 ± 0.096	0.707 ± 0.173	0.066 ± 0.032	0.348 ± 0.071	0.138 ± 0.010	0.817 ± 0.059	0.068 ± 0.019
GAT+HST	0.965 ± 0.004	0.975 ± 0.003	0.688 ± 0.296	0.521 ± 0.289	0.839 ± 0.072	0.138 ± 0.058	**0.974 ± 0.024**	**0.734 ± 0.140**	0.654 ± 0.009	0.034 ± 0.005	**0.914 ± 0.076**	**0.592 ± 0.194**	0.791 ± 0.002	0.037 ± 0.001
GraphSage+RRCF	**0.947 ± 0.041**	**0.955 ± 0.035**	0.811 ± 0.059	0.538 ± 0.069	0.995 ± 0.001	0.679 ± 0.011	0.616 ± 0.097	0.347 ± 0.059	0.804 ± 0.064	0.072 ± 0.025	0.728 ± 0.041	0.226 ± 0.036	0.506 ± 0.105	0.025 ± 0.005
ARES	**0.985 ± 0.008**	**0.991 ± 0.005**	**0.985 ± 0.002**	**0.892 ± 0.014**	**0.999 ± 0.000**	**0.942 ± 0.007**	**0.984 ± 0.001**	**0.806 ± 0.025**	**0.969 ± 0.004**	**0.291 ± 0.014**	**0.956 ± 0.019**	**0.712 ± 0.086**	**0.948 ± 0.049**	**0.249 ± 0.091**

Table 2: Ablation study results on multiple datasets, highlighting the effectiveness of combining GraphSAGE embeddings with Half-Space Trees.

Dataset	ARES	RRCF
DARPA	7.8	33.9
UNSW-NB15	5.4	10.7
ISCX2012	14.1	28.3
CIC-IDS2017	51.4	457.6
CTU-13 Scenario 1	59.5	260.2
CTU-13 Scenario 10	44.9	85.4
CTU-13 Scenario 13	108.3	203.5

Table 3: Average inference time (in seconds) across multiple datasets for ARES **(HST-based) and RRCF.**