Supplementary Materials

Table S1 List of the 38 shared features used for cross-dataset alignment

No.	Shared Feature
1	ARP
2	AVG
3	DHCP
4	DNS
5	HTTP
6	HTTPS
7	Header_Length
8	IAT
9	ICMP
10	IGMP
11	IPv
12	IRC
13	LLC
14	Max
15	Min
16	Number
17	Protocol Type
18	Rate
19	SMTP
20	SSH
21	Std
22	TCP
23	Telnet
24	Tot size
25	Tot sum
26	UDP
27	Variance
28	ack_count
29	ack_flag_number
30	cwr_flag_number
31	ece_flag_number
32	fin_count
33	fin_flag_number
34	psh_flag_number
35	rst_count
36	rst_flag_number
37	syn_count
38	syn_flag_number

Table S2 Dataset split statistics used in the experiments

Task/Dataset	Split	File	Samples	Label Column
CICIoT2023 Binary	Train	datasets\processed\ciciot2023_binary\train.csv	420000	label_id
CICIoT2023 Binary	Validation	datasets\processed\ciciot2023_binary\val.csv	60000	label_id
CICIoT2023 Binary	Test	datasets\processed\ciciot2023_binary\test.csv	120000	label_id
CICIoMT2024 Binary	Train	datasets\processed\ciciomt2024_binary\train.csv	6082355	label_id
CICIoMT2024 Binary	Validation	datasets\processed\ciciomt2024_binary\val.csv	1073357	label_id
CICIoMT2024 Binary	Test	datasets\processed\ciciomt2024_binary\test.csv	1612117	label_id
CICIoT2023 Multi-class	Train	datasets\processed\ciciot2023_7class\train.csv	341348	label_id
CICIoT2023 Multi-class	Validation	datasets\processed\ciciot2023_7class\val.csv	48765	label_id
CICIoT2023 Multi-class	Test	datasets\processed\ciciot2023_7class\test.csv	97529	label_id
Cross-Dataset Aligned Binary	Main Train Aligned	datasets\processed\cross_dataset_aligned_binary\main_train_aligned.csv	420000	label_id
Cross-Dataset Aligned Binary	Main Validation Aligned	datasets\processed\cross_dataset_aligned_binary\main_val_aligned.csv	60000	label_id
Cross-Dataset Aligned Binary	Main Test Aligned	datasets\processed\cross_dataset_aligned_binary\main_test_aligned.csv	120000	label_id
Cross-Dataset Aligned Binary	External Test Aligned	datasets\processed\cross_dataset_aligned_binary\external_test_aligned.csv	1612117	label_id

Table S3 Class distribution statistics used in the experiments

Task/Dataset	Split	Label Column	Class/Label	Samples
CICIoT2023 Binary	Train	label_id	0	210000
CICIoT2023 Binary	Train	label_id	1	210000
CICIoT2023 Binary	Validation	label_id	0	30000
CICIoT2023 Binary	Validation	label_id	1	30000
CICIoT2023 Binary	Test	label_id	0	60000
CICIoT2023 Binary	Test	label_id	1	60000
CICIoMT2024 Binary	Train	label_id	0	5918533
CICIoMT2024 Binary	Train	label_id	1	163822
CICIoMT2024 Binary	Validation	label_id	0	1044447
CICIoMT2024 Binary	Validation	label_id	1	28910
CICIoMT2024 Binary	Test	label_id	0	1574510
CICIoMT2024 Binary	Test	label_id	1	37607
CICIoT2023 Multi-class	Train	label_id	0	56000
CICIoT2023 Multi-class	Train	label_id	1	9144
CICIoT2023 Multi-class	Train	label_id	2	56000
CICIoT2023 Multi-class	Train	label_id	3	56000
CICIoT2023 Multi-class	Train	label_id	4	56000
CICIoT2023 Multi-class	Train	label_id	5	46383
CICIoT2023 Multi-class	Train	label_id	6	56000
CICIoT2023 Multi-class	Train	label_id	7	5821
CICIoT2023 Multi-class	Validation	label_id	0	8000
CICIoT2023 Multi-class	Validation	label_id	1	1307
CICIoT2023 Multi-class	Validation	label_id	2	8000
CICIoT2023 Multi-class	Validation	label_id	3	8000
CICIoT2023 Multi-class	Validation	label_id	4	8000
CICIoT2023 Multi-class	Validation	label_id	5	6626
CICIoT2023 Multi-class	Validation	label_id	6	8000
CICIoT2023 Multi-class	Validation	label_id	7	832
CICIoT2023 Multi-class	Test	label_id	0	16000
CICIoT2023 Multi-class	Test	label_id	1	2613

CICIoT2023 Multi-class	Test	label_id	2	16000
CICIoT2023 Multi-class	Test	label_id	3	16000
CICIoT2023 Multi-class	Test	label_id	4	16000
CICIoT2023 Multi-class	Test	label_id	5	13253
CICIoT2023 Multi-class	Test	label_id	6	16000
CICIoT2023 Multi-class	Test	label_id	7	1663
Cross-Dataset Aligned Binary	Main Train Aligned	label_id	0	210000
Cross-Dataset Aligned Binary	Main Train Aligned	label_id	1	210000
Cross-Dataset Aligned Binary	Main Validation Aligned	label_id	0	30000
Cross-Dataset Aligned Binary	Main Validation Aligned	label_id	1	30000
Cross-Dataset Aligned Binary	Main Test Aligned	label_id	0	60000
Cross-Dataset Aligned Binary	Main Test Aligned	label_id	1	60000
Cross-Dataset Aligned Binary	External Test Aligned	label_id	0	1574510
Cross-Dataset Aligned Binary	External Test Aligned	label_id	1	37607