APPENDIX

Table 1. Chronological Distribution of the Final Dataset

Dataset partition	Benign	DoS	DDoS	Port scanning	Connection flooding	SSH brute force	Total
Training	126,000	31,500	44,100	25,200	18,900	6,300	252,000
Temporal validation	42,000	10,500	14,700	8,400	6,300	2,100	84,000
Streaming test	42,000	10,500	14,700	8,400	6,300	2,100	84,000
Total	210,000	52,500	73,500	42,000	31,500	10,500	420,000

Table 2. Retained Edge-Level Features

No.	Canonical feature	Definition and computation	Type	Preprocessing
1	flow_duration	Elapsed time between the first and final packet associated with a flow, expressed in seconds. A value approaching zero indicates a short or rapidly terminated connection.	Numerical	Log transformation when strongly skewed, followed by standardization
2	forward_packet_count	Number of packets transmitted from the initiating source endpoint to the destination endpoint during the flow.	Numerical count	Log transformation and standardization
3	backward_packet_count	Number of packets transmitted from the destination endpoint back to the initiating source endpoint.	Numerical count	Log transformation and standardization
4	forward_byte_count	Total number of payload-independent packet bytes transmitted in the forward direction. Packet content was not retained.	Numerical count	Log transformation and standardization
5	backward_byte_count	Total number of packet bytes transmitted in the backward direction.	Numerical count	Log transformation and standardization
6	forward_packet_length_mean	Arithmetic mean of packet lengths observed in the forward direction.	Numerical	Standardization
7	forward_packet_length_std	Standard deviation of packet lengths in the forward direction, representing variation in forward packet size.	Numerical	Standardization
8	forward_packet_length_min	Minimum observed packet length in the forward direction.	Numerical	Standardization
9	forward_packet_length_max	Maximum observed packet length in the forward direction.	Numerical	Standardization
10	backward_packet_length_mean	Arithmetic mean of packet lengths observed in the backward direction.	Numerical	Standardization

No.	Canonical feature	Definition and computation	Type	Preprocessing
11	backward_packet_length_std	Standard deviation of packet lengths in the backward direction.	Numerical	Standardization
12	backward_packet_length_min	Minimum observed packet length in the backward direction.	Numerical	Standardization
13	backward_packet_length_max	Maximum observed packet length in the backward direction.	Numerical	Standardization
14	packet_rate	Total number of forward and backward packets divided by flow duration. It represents the overall packet-transmission intensity of a flow.	Numerical rate	Log transformation and standardization
15	byte_rate	Sum of forward and backward bytes divided by flow duration. It measures the total transmission volume per second.	Numerical rate	Log transformation and standardization
16	forward_packet_rate	Number of forward packets divided by flow duration.	Numerical rate	Log transformation and standardization
17	backward_packet_rate	Number of backward packets divided by flow duration.	Numerical rate	Log transformation and standardization
18	flow_iat_mean	Mean inter-arrival time between consecutive packets belonging to the same flow.	Numerical temporal	Log transformation and standardization
19	flow_iat_std	Standard deviation of packet inter-arrival times within a flow. It reflects temporal irregularity or burstiness.	Numerical temporal	Log transformation and standardization
20	flow_iat_min	Minimum observed inter-arrival time between consecutive flow packets.	Numerical temporal	Log transformation and standardization
21	flow_iat_max	Maximum observed inter-arrival time between consecutive flow packets.	Numerical temporal	Log transformation and standardization
22	active_time_mean	Mean duration of active transmission periods in which packet exchanges occurred without exceeding the predefined inactivity threshold.	Numerical temporal	Log transformation and standardization
23	active_time_std	Standard deviation of active-period durations within the flow.	Numerical temporal	Log transformation and standardization
24	idle_time_mean	Mean duration of idle intervals separating active packet-transmission periods.	Numerical temporal	Log transformation and standardization
25	idle_time_std	Standard deviation of idle-period durations within the flow.	Numerical temporal	Log transformation and standardization
26	packet_direction_ratio	Ratio of forward to backward packet counts, calculated using smoothing to prevent division by zero.	Numerical ratio	Clipping and standardization

No.	Canonical feature	Definition and computation	Type	Preprocessing
27	byte_direction_ratio	Ratio of forward to backward byte counts, calculated using smoothing to prevent division by zero.	Numerical ratio	Clipping and standardization
28	transport_protocol	Transport or network protocol associated with the flow, harmonized into TCP, UDP, ICMP, and other protocol categories.	Categorical	One-hot encoding or learned embedding
29	source_port_category	Category of the initiating source port, classified as well-known, registered, dynamic/private, or unavailable.	Categorical	One-hot encoding or learned embedding
30	destination_port_category	Category of the destination port, representing the service range targeted by the flow.	Categorical	One-hot encoding or learned embedding
31	syn_flag_ratio	Number of TCP packets containing the SYN flag divided by the total number of TCP packets in the flow. It represents connection-initiation activity.	Numerical ratio	Zero for non-TCP flows; standardized
32	rst_flag_ratio	Number of TCP packets containing the RST flag divided by the total number of TCP packets in the flow. It represents rejected, interrupted, or forcibly terminated connections.	Numerical ratio	Zero for non-TCP flows; standardized

Table 3. Port-Category Encoding

Category	Port range	General interpretation
Well-known or system	0–1,023	Common system and standardized application services
Registered	1,024–49,151	Registered application and organizational services
Dynamic or private	49,152–65,535	Temporary client-side, private, or dynamically assigned ports
Unavailable	Not recorded or invalid	Missing, unsupported, or non-applicable port information

Table 4. Derived Node-Level Features

No.	Node feature	Definition
N1	incoming_degree	Number of incoming directed communication edges associated with an endpoint during the graph window.
N2	outgoing_degree	Number of outgoing directed communication edges generated by an endpoint during the graph window.
N3	node_total_packets	Total number of packets associated with all incoming and outgoing flows incident on the endpoint.
N4	node_total_bytes	Total number of bytes associated with all incoming and outgoing flows incident on the endpoint.
N5	source_diversity	Number of unique source endpoints that communicated with the node during the graph window.
N6	destination_diversity	Number of unique destination endpoints contacted by the node during the graph window.

No.	Node feature	Definition
N7	mean_node_packet_rate	Mean packet rate across all flows incident on the endpoint.
N8	connection_frequency	Number of flow initiations associated with the endpoint per graph-window duration.

Table 5. Summary of the Feature-Selection Stages

Selection stage	Primary objective	Examples of removed attributes
Identifier removal	Prevent host memorization and privacy disclosure	Raw source IP, destination IP, MAC address, flow identifier
Leakage screening	Prevent direct or indirect use of target information	Attack-generation identifiers, scenario names, post-event fields
Variance filtering	Remove attributes with minimal observed variation	Constant flags, unused protocol fields, empty columns
Missingness assessment	Remove features that could not be consistently reconstructed	Dataset-specific fields with excessive missing values
Correlation analysis	Reduce redundant information and model complexity	Duplicate totals, mathematically equivalent rates, repeated statistics
Importance and stability assessment	Retain theoretically meaningful and stable predictors	Weak or unstable variables with limited validation contribution

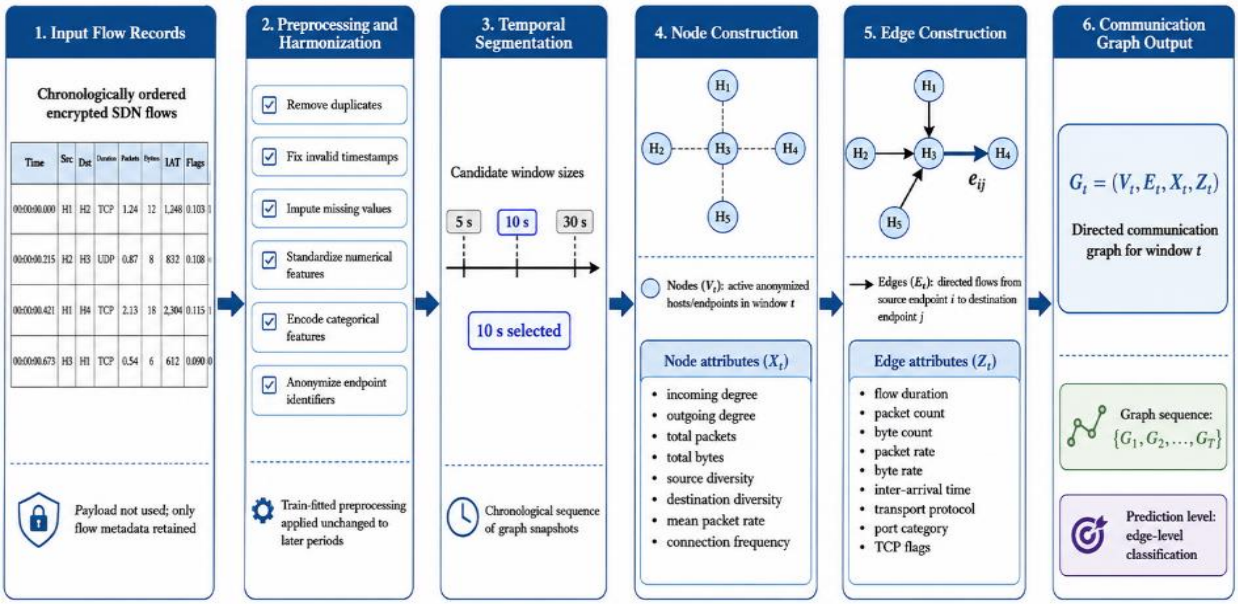

Figure 1. Six-Stage Encrypted SDN Graph Construction Flowchart