Appendix:

A.1 Data Extraction form

Table 1. Data type and item extracted from each study

Data Type	ID	Data Item	Description
Context	D1-D7	Title, author, venue, publication year, publisher, summary, open challenges	Title, author, venue, publication year, publisher, summary including aim, strength, and weakness of the study and Open challenges to be resolved in future.
RQ1	D8-D9	Features, Feature Engineering Method	Features and feature engineering method: automatic or manual used to implement NLP-based HIDS
	D10-D12	Learning Type, Classifier Type, Detection technique	Type of Learning method, Classifier type (e.g., Base, Ensemble), detection technique used for intrusion detection
	D13-D14	HIDS type, Attack detection/classification	Type of HIDS (misuse, anomaly), attack detection (e.g., benign, malicious) or classification (detect specific attack)
RQ2	D15	Attacks	Attacks that are targeted to be detected
RQ3	D16-D17	Data Source, Dataset	Data source or dataset used for training or testing HIDS
RQ4	D18	Evaluation Metric	Metrics used for evaluating HIDS

A.2 Feature types with mapped studies

Table 2. Feature types used in NLP-based HIDS with mapped studies

Feature Type	Study Ref
Statistical (22)	S1, S2, S3, S6, S7, S14, S20, S27, S32, S33, S36, S37, S40, S41, S42, S49, S52, S65, S80, S81, S86, S87
Contextual (49)	S4, S5, S8, S9, S10, S11, S13, S15, S16, S17, S19, S21, S22, S24, S25, S26, S28, S29, S34, S35, S39, S43, S44, S45, S47, S50, S54, S58, S59, S60, S62, S63, S64, S67, S68, S70, S71, S72, S73, S74, S76, S79, S82, S84, S88, S93, S94, S95, S97
Attribute (1)	S18
Temporal (3)	S12, S55, S66
Statistical+Contextual (7)	S46, S61, S77, S78, S85, S96, S98,
Statistical+Attribute (4)	S30, S48, S53, S99
Statistical+Attribute+Temporal (3)	S57, S75, S91
Contextual+Attribute (4)	S23, S31, S51, S56
Temporal+Contextual (3)	S38, S69, S83

Statistical+Contextual+Attribute+Temporal (3)	S89, S90, S92