Appendix A

Table 22: UNSW-NB1515 15 dataset Features based on the description in [5]

No.	Name	Type	Description
1	srcip	nominal	Source IP address
2	sport	integer	Source port number
3	dstip	nominal	Destination IP address
4	dsport	integer	Destination port number
5	proto	nominal	Transaction protocol
6	state	nominal	Indicates to the state and its dependent protocol, e.g. ACC, CLO,, TST, TXD, URH, URN, and (-) (if not used state)
7	dur	Float	Record total duration
8	sbytes	Integer	Source to destination transaction bytes
9	dbytes	Integer	Destination to source transaction bytes

10	sttl	Integer	Source to destination time to live value
11	dttl	Integer	Destination to source time to live value
12	sloss	Integer	Source packets retransmitted or dropped
13	dloss	Integer	Destination packets retransmitted or dropped
14	service	nominal	http, ftp, smtp, ssh, dns, ftp-data ,irc and (-) if not much used service
15	Sload	Float	Source bits per second
16	Dload	Float	Destination bits per second
17	Spkts	integer	Source to destination packet count
18	Dpkts	integer	Destination to source packet count
19	swin	integer	Source TCP window advertisement value
20	dwin	integer	Destination TCP window advertisement value
21	stcpb	integer	Source TCP base sequence number
22	dtcpb	integer	Destination TCP base sequence number
23	smeansz	integer	Mean of the row packet size transmitted by the src
24	dmeansz	integer	Mean of the row packet size transmitted by the dst
25	trans$_d$*epth*	integer	Represents the pipelined depth into the connection of http request/response transaction

26	res$_b dy_l en$	integer	Actual uncompressed content size of the data transferred from the server's http service.
27	Sjit	Float	Source jitter (mSec)
28	Djit	Float	Destination jitter (mSec)
29	Stime	Timestamp	record start time
30	Ltime	Timestamp	record last time
31	Sintpkt	Float	Source interpacket arrival time (mSec)
32	Dintpkt	Float	Destination interpacket arrival time (mSec)
33	tcprtt	Float	TCP connection setup round-trip time, the sum of 'synack' and 'ackdat'.
34	synack	Float	TCP connection setup time, the time between the SYN and the SYNACK packets.
35	ackdat	Float	TCP connection setup time, the time between the SYNACK and the ACK packets.
36	is$_s m_i ps_p orts$	Binary	If source (1) and destination (3) IP addresses equal and port numbers (2)(4) equal then, this variable takes value 1 else 0
37	ct$_s tate_t tl$	Integer	No. for each state
38	ct$_f lw_h ttp_m thd$	Integer	No. of flows that has methods such as Get and Post in http service.

39	is$_{f}tp_{l}ogin$	Binary	If the ftp session is accessed by user and password then 1 else 0.
40	ct$_{f}tp_{c}md$	integer	No of flows that has a command in ftp session.
41	ct$_{s}rv_{s}rc$	integer	No. of connections that contain the same service
42	ct$_{s}rv_{d}st$	integer	No. of connections that contain the same service
43	ct$_{d}st_{l}tm$	integer	No. of connections of the same destination address
44	ct$_{s}rc_{l}tm$	integer	No. of connections of the same source address
45	ct$_{s}rc_{d}port_{l}tm$	integer	No of connections of the same source address
46	ct$_{d}st_{s}port_{l}tm$	integer	No of connections of the same destination address
47	ct$_{d}st_{s}rc_{l}tm$	integer	No of connections of the same source .
48	attack$_{c}at$	nominal	The name of each attack category.
49	Label	binary	0 for normal and 1 for attack records

Table 23: NSL-KDD Attack Type

Major Attacks	Type
DOS	Back
	Land,Neptune
	Pod, Smurf
	Teardrop,Mailbomb
	Processtable,Udpstor
	Apache2, Worm
R2L	Guesspassword
	Ftpwrite
	Imap,Phf
	Multihop,Warezmaster
	Xlock,Xsnoop
	Snmpguess,Snmpgetattack
	Httptunnel,Sendmail
	Named
U2L	Bufferoverflow
	Loadmodule
	Rootkit
	Perl
	Sqlattack
	Xterm,Ps
Probe	Satan
	IPsweep, Nmap
	Portsweep,Mscan
	Sa,int

Appendix B

B.1 Results

Table 24: Accuracy Rate Results of adversarial attacks in ANN. The adversarial attack methods row indicates the methods used that generate adversarial samples by the min-max method

Dataset	Model	Natural	FGSM	CW	BIM	PGD	Deepfool
UNSWNB	ANN	96.64	53	71.62	43.27	43.27	37.87
NSLKDD	ANN	96.07	81.73	19.22	30.49	31.52	29.60

Table 25: Accuracy Rate Results of Adversarial Training for ANN Experiments using min-max. The adversarial attack methods row indicates the methods used that generate adversarial samples by the min-max method

Dataset	Model	Natural	FGSM	CW	BIM	PGD	Deepfool
UNSWNB	ANN	96.64	95.95	90.83	96.21	96.21	94.57
NSLKDD	ANN	96.07	95.99	95.59	95.87	94.61	95.32

Table 26: Accuracy Rate Results of adversarial attacks in CNN. The adversarial attack methods row indicates the methods used that generate adversarial samples by the min-max method

Dataset	Model	Natural	FGSM	CW	BIM	PGD	Deepfool
UNSWNB	CNN	96.19	71.31	66.20	46.52	46.52	66.94
NSLKDD	CNN	95.69	32.48	23.37	49.60	18.71	12.63

Table 27: Accuracy Rate Results of Adversarial Training for CNN Experiments using min-max. The adversarial attack methods row indicates the methods used that generate adversarial samples by the min-max method

Dataset	Model	Natural	FGSM	CW	BIM	PGD	Deepfool
UNSWNB	CNN	96.19	94.51	86.98	95.66	95.47	92.05
NSLKDD	CNN	95.69	95.69	92.67	95.69	95.52	94.31

Table 28: Accuracy Rate Results of adversarial attacks in RNN. The adversarial attack methods row indicates the methods used that generate adversarial samples by the min-max method

Dataset	Model	Natural	FGSM	CW	BIM	PGD	Deepfool
UNSWNB	RNN	95.47	46.92	5.76	18.71	18.71	31.59
NSLKDD	RNN	95.69	44.74	43.94	43.15	43.15	10.36

Table 29: Accuracy Rate Results of Adversarial Training for RNN Experiments using min-max. The adversarial attack methods row indicates the methods used that generate adversarial samples by the min-max method

Dataset	Model	Natural	FGSM	CW	BIM	PGD	Deepfool
UNSWNB	RNN	96.19	94.51	86.98	95.66	95.47	92.05
NSLKDD	RNN	95.69	95.69	92.67	95.69	95.52	94.31