Security in AI

By: Ethan Lung, Damon Lin, Jacob Morin, Rut Mehta

Team Members



Jacob Morin (UG)



Damon Lin (UG)



Ethan Lung (HS)



Rut Mehta (UG)









Honglu Li (Mentor)

Prof. Chen (Advisor) Tianfang Zhang (Mentor) Changming Li (Mentor)

Overview and Goal

- Study the Security of WiFi sensing systems under adversarial attack
- Utilize mobile device to extract channel state information (CSI) to train deep learning model for recognition tasks
 - Human Activity Recognition and User Authentication
- Develop a type of adversarial attack algorithm to generate perturbation that can deceive deep learning model



Motivation

- New WiFi sensing techniques use a learning-based approach for accurate and efficient recognition tasks
 - Activity Recognition and User Authentication
 - Applied to various commodity devices (i.e., smart phone, laptop, routers)
 - Passive engagement from legitimate user
- Learn-based techniques (i.e., Deep learning) are vulnerable against adversarial attack
 - Input data can be easily manipulated
 - Model can be fooled by the attacker to cause false recognition results
- Research in adversarial attacks against WiFi sensing system can reveal security issues
 - Sensing system can be targeted and result serious privacy and security concerns
 - For instance, legitimate user is blocked by his/her property due to the attack

IoT Devices

Internet of Things (IoT) Devices

- Connected to internet
- Gather information
- Can pick up interference from
 - behavioral characteristics
 - body movements
- Send and receive data through wireless channels
- Channel State Information (CSI) has the communications between devices





Deep Learning Models

- Deep Learning Models possess the capability to extract features from input data (individuals)
- These models are able to categorize the input data into specific labels, such as daily activities
 - Walking, Kicking, Raising Arm, Squatting, Sitting



Setting Up Experiments

- Setting up a Linux virtual machine through VirtualBox
 - Ubuntu ISO (disk image file)
 - Increased familiarity with Linux terminal
- Setting up phones on Virtual machine
- Preparing to collect CSI data

The C-based Firmware Patching Framework Our software may damage your hardware and may void your hardware's warranty! You use our tools at your own risk and responsibility read /home/sudouser/nexmon/STATISTICS.md for more infor mation obj, gen, log src/version.c => obj/version.o (details: log/compiler.log) gen/nexmon.pre => gen/nexmon2.pre gen/nexmon.pre => gen/nexmon.ld gen/nexmon.pre => gen/flashpatches.ld => gen/patch.elf (details: log/linker.log, log/linker.err) gen/nexmon.pre => gen/nexmon.mk gen/nexmon.pre => gen/flashpatches.mk gen/flashpatches.mk => fw bcmdhd.bin (details: log/flas hpatches.log) gen/nexmon.mk => fw_bcmdhd.bin (details: log/patches.log) fw bcmdhd.bin => /sdcard/fw bcmdhd.bin /sdcard/fw_bcmdhd.bin => /vendor/firmware/fw_bcmdhd.bin





Collecting Data

- Used WiFi Transmitters/Receivers to collect CSI data
- Performed daily movements (walking, squatting, etc)
- Data was then trained using deep learning model

E	e Edit View Ge	Capture Analyzo Flat	1000			
1			usitics Teleph	ony Wireless Tools	P state of the second s	
				e 🛪 👲 🗖 🖪	QQQU	
: <u>u</u>	wian.so == 82:7b:b	e:f0:98:e0			1 C E	
No.	Time	* Source	PCCI			
- 220	291 2 1496048	82:7b:be:r0:95:e0	-53 dBn	Destination	Protocol Len Info	
1000	293 2.161293	82:7b:be:re:95:e0	-55 dBn	Broadcast	802.11 3. Beacon frame, SA.748 CM.9 Flam	
1000	294 2.162702	82:7h:he:F0:06:00	-56 dBn	Broadcast	002.11 3. Beacon Franc, SN:758, FN:8, Flants	R8, SSID: HyCovertChannel
	295 2.165699	82:7b:be:f0:96:e0	-58 dBa	Broadcast	892.11 3. Beaton frame St. Tes, Flags C, ST-5	98. SSID: McCourt (Channel
1933	297 2.170053	82:7b:be:f0:96:e9	-57 dBn	Broadcast	592.11 3. Baacon frame, SN-758, FM-8, FlagsC, 81-1	09, SSID-MyCovertChangel
1823	298 2.173951	82:70:00:00:00:00:00	-58 dBn	Broadcast	802.11 J. Beacen frame, SN=758, FN=8, Flats- C. 81-1	00, SSID-MyCovertChuntel
1000	299 2.175457	82:7b:be:f6:96:e0	-De dBa	Broadcast	802.11 3. Bracon frame, SN-758, FN-8, FlagsC, BI-11	99, SSID-MyCover(Channel
	391 2.257277 302 3 250475	82:7b:be:f8:96:e0	-69 dBm	Broadcast	802.11 3. Beacon frame, SN-758, FN-8, FlagsC, 81=50	10, SSID=MyCovertChannel
1000	383 2.263725	82:70:be:f0:96:e0	-58 dBm	Broadcast	507 11 3. Beecon frame, SN=758, FM=9, FlagsC. Blats	19. SSID-MyCovertChantel
- 192	384 2.266271	82:70:0e:10:95:e9	-59 cl8m	Browdcast	802.11 3. Beacon frame, SW=758, FN=0, Flags=C, BI=16	9. SSID:HyCovertChannel
. 1999	385 2.387493	82:7b:be:f0:95:e0	-50 088	Broadcast	802.11 3. Bescon frame, SNu758 EN-D Classe	00, SSID-MyCovertChannel
1253	305 2.310183	82:7b:be:f0:95:e0	-42 dBm	Broadcast	802.11 3. Seacon frame, SM:758, FN=0, FLags C. 81-10	6, SSID:MyCovertChannel
-	208 2 312074	8217b:be:f0:96:40	-41 dBm	Broadcast	592.11 8. Seacco frame, SN-758, FN-8, Flags=C, 81-10	10, SSID-PrOvertChannel
1000	309 2.316479	82:7b:be:f8:96.e0	-41 d8n	Broadcast	392.11 3. Boloon Frame, Survey, 51-9, 51agarC, BI-18	H9, SSID:MyCovertChannel
. 233	310 2.321845	82:7b:be:f0:96:e0	-41 (Ba	Broadcast	802.11 3. Beacon Frane, SN+758, FN+0, Flags=	B. SSIDNY(OvartCharne) SSIDNY(OvartCharne)
2000	311 2.361546	82:7b:be:f0:96:e9	-49 dBa	Broadcast	892.11 3. Seacon Frame, SM=758, FN:0, Flags=C, 81=10	9, SSID-MyCovertChantel
1000	313 2.370074	82:70:0e:00:05:00	-48 dBa	Broadcast	802.11 3. Beacon frame, SAT758, FNeB Flags	0, SSID:#yCovertChansel
- 088	314 2.373868	82:7b:be:f0:96:e0	-48 dBm	Broadcast	802.11 3. Beacon frame, SN=758, FN=8, Flags=C. BI=10	9. SSID: MyCover(Charmal
1 1000	315 2.433871	82:7b:be:f0:96:e0	-51 dBm	Broadcast	BRZ 11 3 Bearon frame, SN=758, FN=0, Flags=C, BI=184	0, SSID=MyCovertChannel
1 200	317 2 440022	82:7b:be:f6:96:e0	-51 d8m	Broadcast	892.11 3. Beacon frame, SM+758, FN29, Flams- C BI-100	P, SSID-MyCovertChannel
- 200	318 2.446945	82:7b:be:f8:95:c8	-52 088	Broadcast	802.11 3. Beacon frame, SW=758, FM=0, Flags=C, BI=100	3. SSID-MyCovertCharpel
1000	319 2.444124	82:7b:be:f0:95:c0	-S1 dBn	Broadcast	802.11 3. Beacon frame, SN=758, FN=0, Flags=C, BI=100), SSID:HyCovertChannel
1000	328 2.458274	82:7b:be:f0:95:e0	-53 dBn	Broadcast	#02.11 3. Beacon frame, Shu758 DN=0 Claure C BI-100	3, SSID-MyCovertChannel
4					And is a former during the state of the state of the state	Anto a A
1000	Signal strength	(dBm): -41dBm				
	(Duration: 444)	19960760				
💌 1E	EE 002.11 Beacor	frame, Flags:	c			
1000	Type/Subtype: 8	eacon frame (8x0008)	Sales Columb	States and the second states of	The second se	
	.000 0001 P01 P	9998 = Duration: 0 pior	oseconts			
1000	Receiver addres	s: Broadcast (ff:ff:ff:				
1000	Destination add	ress: Broadcast (ff:ff:	fforforforf)			
1000		ress: 82:70:be:f8:96:e8		9195140)		
	BSS Id: 82:7h:b	e:f8:96:e0 (82:7b:be:f9	195:001			
		0000 = Fragment number:	0			
0.33	8818 1111 0110	= Sequence number:	758			
1000	IFCS Status: 60					
1 70	EE Sta 11 wirele	SS LAN				
P . 15						
1						
6020	60 00 d3 00 d7	01 10 00 00 00 ff ff ft				



Results

- Used a Confusion Matrix to check accuracy
- Model is able to achieve recognition accuracy at 96% for User authentication
- Overall attack success rate can reach to 80%



