

פרויקט גמר - רשתות

ירין לוי - 213907421
מאיה חייט - 322515669

קישור לgithub :

<https://github.com/MayaHayat/NetworkingFinalProject.git>

קישור ל LinkedIn :

Maya:

https://www.linkedin.com/posts/maya-hayat-855777288_github-mayahayatnetworkingfinalproject-activity-7097143783804293120-YIn5?utm_source=share&utm_medium=member_desktop

Yarin:

https://www.linkedin.com/posts/yarin-levi-a7bb97288_github-thelegitbaljitnetworkingfinalproject-activity-7097519351271976960-dLkq?utm_source=share&utm_medium=member_desktop

תוכן עניינים

2	חלק יבש
4	חלק רטוב
4	מבוא
4	מהלך העבודה
8	הקלטה 1: הודעות רגילות
9	הקלטה 2: תמונות
11	הקלטה 3: הודעות קוליות
12	נקודה חשובה
13	הקלטה 4: סרטונים
14	הקלטה 5: משלבים הכל ביחד
15	הקלטה 6: הוספת רעש להקלטות
16	ניתוח התעבורה אצל שני הצדדים
17	סיכום

חלק יבש

כיום בעולם, ישנם יותר מ-2 מיליארד משתמשים באפליקציות לשליחת הודעות (IM). כתוצאה מכך ממשלות ותאגידים גדולים מעוניינים לנטר את המשתמשים הללו ברשת על מנת להשפיע על תוכן שהאזרחים צורכים ואף לנטר איומים כלפי הממשלה. המאמר "Practical Traffic Analysis Attacks on Secure Messaging Applications" מדבר על ניתוח תנועת שליחת הודעות באינטרנט ומדגיש את הקלות שבה ניתן לזהות את המשתתפים והמנהלים בקבוצות מסוימות, על אף ההצפנה הנוקשה של אפליקציות אלה (לדוגמא: end to end encryption, end to middle encryption) הכותב אף מציין כי 15 דקות של שיחה ב-Telegram מספיקות על מנת לזהות את המשתמשים והמנהלים בדיוק של 94% עם שגיאה נמוכה ביותר.

התוקף יכול לנטר את המידע המבוקש על תנועות השיחה ע"י ניתוח המידע, כלומר הוא יאסוף מידע על תזמון שליחת הודעות, אורך ההודעות ומספר ההודעות הנשלחות בחלון זמן מסוים על מנת לזהות את המנהל ומשתתפי הקבוצה (את כתובות ה-IP שלהם) ובכך יכולה הממשלה או כל גוף אחר המעוניין במידור השיחות לשלוט במידע המועבר בקבוצות. ישנן 3 דרכים למצוא את האינפורמציה הרצויה של ערוץ: הראשונה כי כמובן אם השיחה (ערוץ) לא פרטית (public group chat) התוקף יכול להצטרף בתור משתתף ולהקליט את הפרטים הנדרשים (זמן וגודל ההודעות כפי שצינו לעיל). השנייה היא שהתוקף הצטרף לקבוצה ויכול לשלוח הודעות (או שהקבוצה מאפשרת לכולם לפרסם הודעות או שהוא קיבל אישור מיוחד - גישת מנהל), מה שעוזר לו לשלוח כעת הודעות בעצמו ולנטר את הפרטים שהוא מעוניין בהם על ידי שליחת ההודעות בקצב הנדרש ובאורך מסוים שיכול לעזור לו. האופציה השלישית והאחרונה היא שהתוקף לא מצליח להיכנס לקבוצה (channel) אבל הוא כן הצליח לגלות את הכתובת IP של אחד המשתתפים או המנהלים בקבוצה, התוקף מאזין לתנועה המוצפנת של המשתתף שמצא ומקליט את הפרטים.

התוקף מבצע האזנת סתר (Wiretapping) באמצעות ניטור התעבורה של המידע המוצפן על מנת לזהות את המשתמשים והמנהלים בקבוצה, הדבר מתבצע באמצעות האזנת סתר של תעבורת הרשת של ספקי האינטרנט או ה-IXP שבהם הוא שולט. למשל, באמצעות חומת האש הגדולה של סין (Great Firewall Of China) לחלופין, היריב יכול להאזין לאנשים ספציפיים (למשל, פעילים חשודים).

שירותי IM מאפשרים למשתמשים שלהם לשלוח סוגים שונים של הודעות, לרוב טקסט, תמונה, וידאו, קובץ אודיו והודעות. הודעות IM מועברות בין משתמשים דרך אחת מצורות התקשורת העיקריות הבאות:

- תקשורת אחד על אחד בין משתמשי IM מכונה "הודעות ישירות". בשירותי IM פופולריים מרכזיים, הודעות ישירות אלו מועברות דרך השרתים של ספקי ה-IM. אלא אם כן ישנה הצפנה מקצה לקצה, לשרתים יש גישה לתוכן התקשורת הללו.
- לאחר מכן, יש לנו "תקשורת קבוצתית פרטית (סגורה)" שבה משתמשים מרובים עוסקים בתקשורת. בתוך קבוצות אלה, כל החברים יכולים לפרסם הודעות ולקרוא את מה שאחרים

פרסמו. לכל קבוצה יש מנהל שיצר אותה ובעל יכולת לנהל משתמשים והודעות. כדי להצטרף לקבוצה סגורה, משתמשים צריכים לקבל הזמנה.

- לבסוף, יש לנו "תקשורת קבוצתית ציבורית" (פתוחה) הידועה גם בשם "ערוצים". צורת תקשורת זו פועלת כשידור, שבו מנהל מערכת אחד או יותר יכולים לפרסם הודעות, וחברים יכולים לקרוא רק פוסטים אלה. הצטרפות לערוצים אינה מצריכה הזמנה; משתמשים יכולים להצטרף באופן חופשי.

מהטבלה ניתן להסיק כי המרכיב המכריע בסוג ההודעות הינו תמונות, המרכיבות 48% מכלל ההודעות, וגודלן הממוצע הוא כ-306.61 בתים. לאחר מכן הודעות טקסט, המרכיבות כ-29.4% מכלל ההודעות וגודלן הממוצע הוא כ-91.33 אלף בתים. ולאחר מכן קבצי וידאו (15.4%), אודיו (5.1%) וקבצים (2.1%).

כותב המאמר מציין כי פיתחו שני אלגוריתמים לניתוח התקפות SIM, כפי שצינו מקודם מטרת ההתקפות היא לזהות את מנהלי הקבוצות והמשתתפים, מה שמאפשר לנו את ההתקפה היא העובדה שאין טשטוש של מאפייני הקבוצה (כלומר אנחנו יודעים את תזמון ההודעות, גודלן והתדירות בה הן נשלחות. האלגוריתם הראשון הוא "event based correlator" שמשתמש בנתונים סטטיסטיים של הקבוצות שמתאים משתמשים לערוצים. האלגוריתם השני הוא "shaped based algorithm" שכמו השם שלו מתאים את הצורות של התנועה על מנת להתאים את המשתמשים לערוצים/קבוצות.

תפקיד התוקף הוא להחליט בין שני מקרים, האם משתמש U נמצא בערוץ C, כלומר הוא פעיל ושולח הודעות בתור משתתף או בתור מנהל או שהמקרה השני שהמשתמש לא קשור לערוץ C. כיוון שהערוץ מוצפן התוקף משתמש הגלאי (event based detector) מעוניין להתאים בין אירועי SIM של משתמש U לאירועים של ערוץ C, כאשר אירוע E יכול להיות הודעת SIM אחת או מספר הודעות שנשלחו בחלון זמן מוגבל. אנחנו מסתכלים על אירוע כאחד מהחמישה: תמונה, סרטון, קובץ, הודעת טקסט או הודעה קולית. כאשר אנחנו מסתכלים על אירוע מסוים אנחנו מקליטים את הגודל S שלו ואת הזמן T בו הוא הופיע. תמונה 8 ניתן לראות קווים קטנים ודקים שנוצרו על ידי הקלטת התראות וכו', כיוון שהם חסרי משמעות הגלאי מתעלם מהם.

תמונה 8 במאמר מציגה את חלוץ האירועים שהתקבלו על ידי שני משתמשים U1 העליון וU2 התחתון. אנחנו רואים שאין התאמה בין תזמון שליחת הסרטון לתנועה שמתבצעת אצל U1 (אין תנועה בזמן שליחת הסרטון) בעוד שיש התאמה מדויקת בין תזמון ההודעות לתנועה של U2 לכן אנחנו יכולים להסיק שU2 אכן קשור לקבוצה C. משהו שחשוב לשים לב אליו זה אורך כל הודעה, כמובן שסרטון יתפוס יותר זמן מאשר תמונה או קובץ טקסט, מידע זה יכול לעזור לנו למנוע טעויות.

חלק רטוב

מבוא

בחלק הרטוב התבקשנו להקליט תקשורת ב- 4 קבוצות IM שונות, בפרויקט שלנו השתמשנו WhatsApp Web ובWireshark על מנת להקליט את השיחות כלומר את הפקטות שנשלחו והתקבלו. השתמשנו בפייטון בשביל ליצור גרפים של גודל פקטות כציר ה Y וזמן ההקלטה כציר ה X ובכך נוכל בהמשך להראות בדיוק איזה סוג הודעה התקבלה ומי שלח איזה הודעה ובעזרת הקלטות וניתוח זה יוכל ה"תוקף" לקבל את המידע הנחוץ כלומר בהינתן העובדה שהתוקף כבר נמצא בקבוצה (שהקבוצה הייתה פתוחה ולכן יכול היה להצטרף או שמישהו שכבר היה חבר בקבוצה צירף אותו) יוכל התוקף לזהות מתי נשלחו הודעות ומי נמצא בקבוצה הנוכחית.

חילקנו את העבודה ל 5 הקלטות שונות בכל פעם הקלטנו קבוצה שונה (קבוצת הודעות טקסט, תמונות, הודעות קוליות ושיחה אחת של כולם ביחד וניסינו לזהות מה נשלח בכל הודעה) להקלטה החמישית הוספנו רעש שיצרנו..

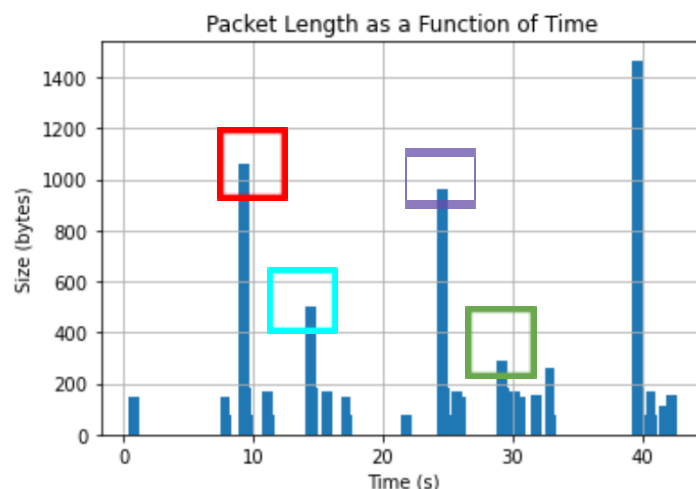
מהלך העבודה

תחילה נשים לב כי :

```
C:\Users\Maya>nslookup web.whatsapp.com
Server: UnKnown
Address: 2a0d:6fc2:5560:c200::1

Non-authoritative answer:
Name: mmx-ds.cdn.whatsapp.net
Addresses: 2a03:2880:f276:cd:face:b00c:0:167
157.240.251.60
Aliases: web.whatsapp.com
```

כלומר או שהsource או הdestination של הפקטות אמורות להיות הכתובת שמסומנת מעל, כפי שניתן לראות בהקלטה הבאה זה אכן המצב:



ניתן לראות בבירור את הקפיצה בגובה קצת מעל 1,000 סביב ה 10 שניות ומיד אחרי זה קפיצה של 500 כ 5 שניות לאחר מכן ועוד קפיצה בגובה קצת פחות מהקפיצה הראשונה סביבות ה 25 שניות. בתמונה מעל סימנו את הקפיצות והפאקטות המתאימות בכל פעם בצבע שונה.

No.	Time	Source	Destination	Protocol	Length	Info
146	7.877545	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=1 Ack=141 Win=1141 Len=0
207	9.292341	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	1059	Application Data
208	9.361159	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=1 Ack=1126 Win=1152 Len=0
209	9.462018	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TLSv1.2	185	Application Data
210	9.516298	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TCP	74	52750 → 443 [ACK] Seq=1126 Ack=112 Win=514 Len=0
212	11.046223	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	144	Application Data
213	11.114447	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=112 Ack=1196 Win=1152 Len=0
214	11.123377	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TLSv1.2	168	Application Data
215	11.126300	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	167	Application Data
216	11.237726	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=206 Ack=1280 Win=1152 Len=0
221	14.380524	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	499	Application Data
222	14.440440	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=206 Ack=1714 Win=1163 Len=0
225	14.551144	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TLSv1.2	185	Application Data
226	14.599955	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TCP	74	52750 → 443 [ACK] Seq=1714 Ack=317 Win=513 Len=0
241	15.654545	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TLSv1.2	168	Application Data
242	15.657258	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	167	Application Data
243	15.726272	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=411 Ack=1807 Win=1163 Len=0
252	17.129797	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	144	Application Data
253	17.198215	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=411 Ack=1877 Win=1163 Len=0
271	21.728191	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TCP	75	55177 → 443 [ACK] Seq=1 Ack=1 Win=511 Len=1 [TCP segment of a reassembled PDU]
272	21.811918	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 55177 [ACK] Seq=1 Ack=2 Win=278 Len=0
278	24.551597	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	144	Application Data
279	24.619524	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=411 Ack=1947 Win=1163 Len=0
280	24.633957	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	963	Application Data
281	24.702932	2a03:2880:f276:cd:f...	2a0d:6fc2:5560:c200...	TCP	74	443 → 52750 [ACK] Seq=411 Ack=2836 Win=1174 Len=0

בשביל לבדוק כי ההקלטות שלנו אמינות תזמנו את ההקלטות וכתבנו בהודעות את הזמנים שבהם ההודעות נשלחו כמו הדוגמא הבאה (כמובן שהצילומי מסך אינם ההודעה המלאה כפי שיכולנו לראות מהגרף וצילומי המסך של הפקטות כל הודעה הייתה באורך שונה.

10	import pandas as pd import matplotlib.pyplot as plt import pandas as pd import matplotlib.pyplot as plt
15	import pandas as pd import matplotlib.pyplot as plt import pandas as pd import matplotlib.pyplot as plt
25	import pandas as pd import matplotlib.pyplot as plt import pandas as pd import matplotlib.pyplot as plt import pandas as pd

כפי שניתן לראות בצילומי מסך ההודעות שנשלחו הופיעו בדיוק באותו הזמן בגרף כפי שנשלחו לכן נוכל לסמוך על ההקלטות שלנו מכאן הלאה. כמובן שיכול להראות מוזר שבחרנו הודעות ארוכות בשביל ליצור קפיצות גבוהות בגרף לכן שלחנו גם את ההודעה הבאה וכפי שניתן לראות יש קפיצה גם סביב ה 30 שניות (מסומן בירוק).

30 היייווי
14:19 ✓

לאחר שפתחנו 4 קבוצות וואצטאפ שונות, החלטנו לפני כל הקלטה על סוגי ההודעות שאנחנו מעוניינים לשלוח ותזמונים להודעות. הקלטנו בכל פעם במשך 90 שניות וסיננו לפי התמונה הבאה:

(tcp.port == 443 tcp.port == 5222) && (ipv6.src == 2a03:2880:f276:cd:face:b00c:0:167 ipv6.dst == 2a03:2880:f276:cd:face:b00c:0:167)						
No.	Time	Source	Destination	Protocol	Length	Info
2	0.737011	2a0d:6fc2:5560:c200...	2a03:2880:f276:cd:f...	TLSv1.2	144	Application Data

אחרי שסיימנו להקליט העברנו את ההקלטה לקובץ csv מהצורה הבאה

	A	B	C	D	E	F	G
1	No.	Time	Source	Destination	Protocol	Length	Info
2	2	0.737911	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TLSv1.2	144	Application Data
3	3	0.807164	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TCP	74	443 > 52750 [ACK] Seq=1 Ack=71 Win=1141 Len=0
4	145	7.809421	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TLSv1.2	144	Application Data
5	146	7.877545	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TCP	74	443 > 52750 [ACK] Seq=1 Ack=141 Win=1141 Len=0
6	207	9.292341	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TLSv1.2	1059	Application Data
7	208	9.361159	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TCP	74	443 > 52750 [ACK] Seq=1 Ack=1126 Win=1152 Len=0
8	209	9.463018	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TLSv1.2	185	Application Data
9	210	9.516298	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TCP	74	52750 > 443 [ACK] Seq=1126 Ack=112 Win=514 Len=0
10	212	11.046223	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TLSv1.2	144	Application Data
11	213	11.114447	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TCP	74	443 > 52750 [ACK] Seq=112 Ack=1196 Win=1152 Len=0
12	214	11.123377	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TLSv1.2	168	Application Data
13	215	11.1263	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TLSv1.2	167	Application Data
14	216	11.237726	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TCP	74	443 > 52750 [ACK] Seq=206 Ack=1289 Win=1152 Len=0
15	221	14.380524	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TLSv1.2	499	Application Data
16	222	14.44844	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TCP	74	443 > 52750 [ACK] Seq=206 Ack=1714 Win=1163 Len=0
17	225	14.551144	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TLSv1.2	185	Application Data
18	226	14.599955	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TCP	74	52750 > 443 [ACK] Seq=1714 Ack=317 Win=513 Len=0
19	241	15.654545	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TLSv1.2	168	Application Data
20	242	15.657258	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TLSv1.2	167	Application Data
21	243	15.726272	2a03:2880:f276:cd:face:b00c:0:167	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	TCP	74	443 > 52750 [ACK] Seq=411 Ack=1807 Win=1163 Len=0
22	252	17.129797	2a0d:6fc2:5560:c200:bdce:1f2e:bc44:2269	2a03:2880:f276:cd:face:b00c:0:167	TLSv1.2	144	Application Data

מהטבלה אנחנו יכולים ללמוד שכמו שאמרנו ההקלטה ארכה 42.3 שניות (המקסימום בזמן) האורך הממוצע של הפקטות הן 177 והחציון הוא 128 (כמובן שה-STD גבוה כיוון שישנן הודעות קצרות וארוכות בהקלטה).

ולאחר מכן השתמשנו בקוד שכתבנו מראש בשביל ליצור את הגרפים שראינו מקודם על מנת שנוכל לנתח את ההודעות היוצאות והנכנסות בקבוצות שהתווקף נמצא בהן.

```
data.describe()
```

	No.	Time	Length
count	64.000000	64.000000	64.000000
mean	274.984375	25.257952	177.281250
std	74.456655	11.863967	237.391477
min	2.000000	0.737911	74.000000
25%	225.750000	14.587752	74.000000
50%	286.500000	25.776032	128.500000
75%	318.250000	32.892591	168.000000
max	377.000000	42.326907	1466.000000

נשים לב כי בהקלטה יש מספר רב של פקטות באורך 74 ביטים (ולפי הטבלה שמעל 25% מהפקטות) שהן מסוג ACK כפי שלמדנו בקורס פקטות מסוג זה אחראיות לשלוח הודעה מהמקבל לשולח כי הוא אכן קיבל את ההודעה לכן נניח בהקלטות אלה שמדובר ברעש. כמובן שפקטות בגודל זה קטנות מדי בשביל להיות הודעה(כיוון שהודעות קטנות (למשל ההודעה ששלחנו בזמן 30) הן סביבות 150 ביטים לפחות).

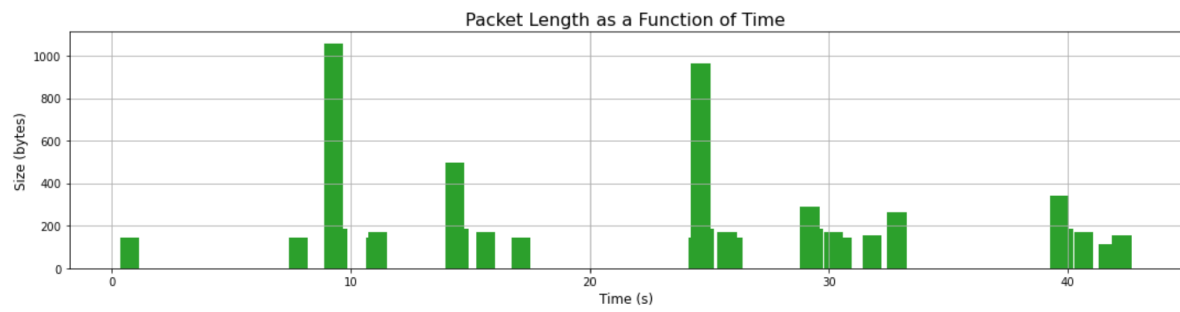
```
filtered_df = data[data['Info'] == 'Application Data']
```

ננסה להוריד את כל השורות המכילות את המילה ACK בעמודת ה Info.

```
filtered_df.describe()
```

נשים לב שהממוצע וגם החציון עלה והרבעון התחתון יושב כרגע על 145 מספר הרבה יותר הגיוני מאשר 74. נראה איך הגרף נראה עכשיו.

	No.	Time	Length
count	32.000000	32.000000	32.000000
mean	269.84375	24.418040	237.031250
std	72.57037	11.637967	215.888036
min	2.000000	0.737911	113.000000
25%	224.000000	14.508489	145.500000
50%	285.500000	25.739643	167.500000
75%	313.500000	32.069422	185.000000
max	376.000000	42.272604	1059.000000

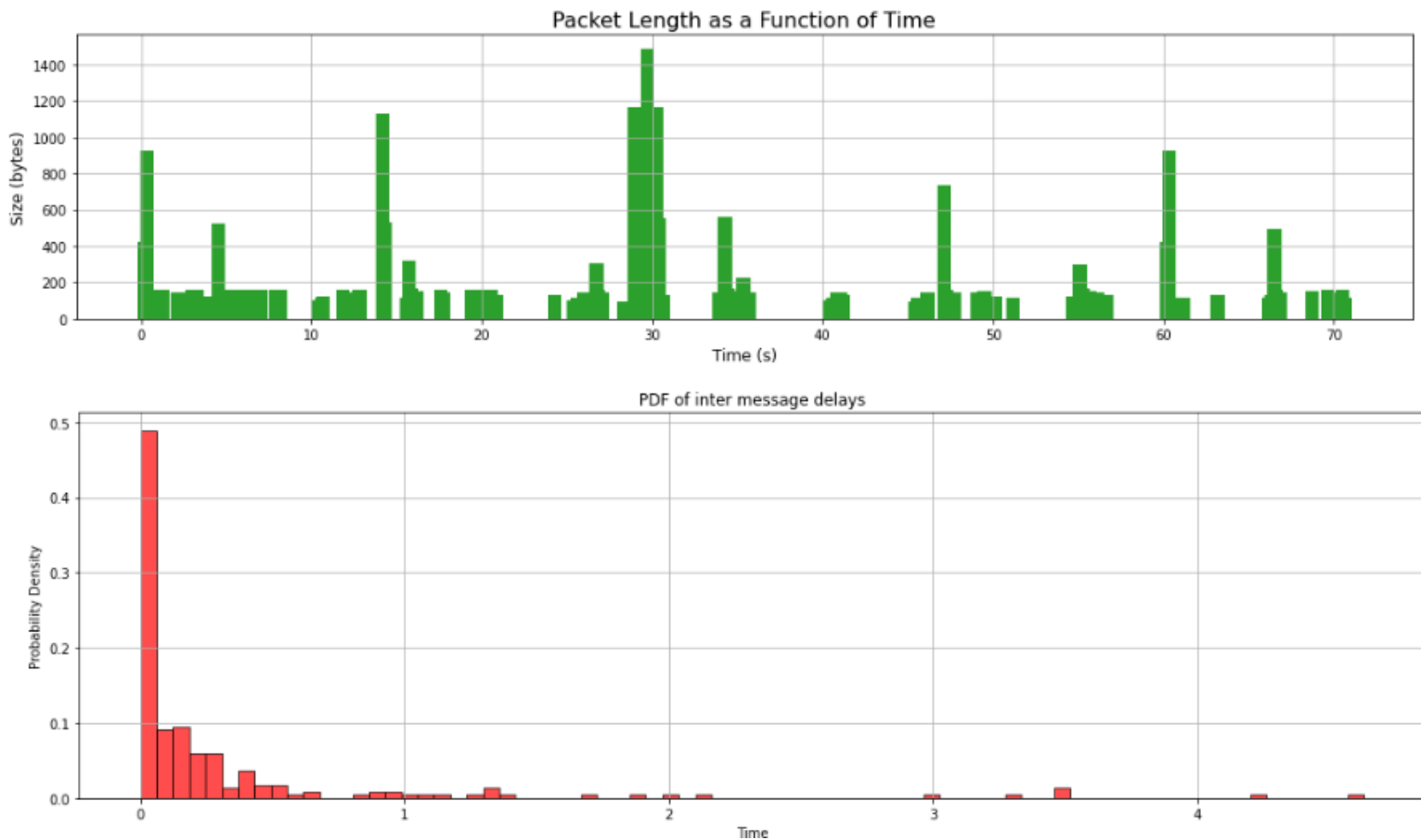


נשים לב שהגרף הרבה יותר נקי עכשיו, הורדנו את כל הקפיצות הקטנות (Acks) כלומר את הרעש וכעת ניתן לראות בדיוק מתי נשלחו איזה הודעות. נשתמש מכאן בקוד המפלט.

הקלטה 1: הודעות רגילות

התוקף מתחיל את ההקלטה ושולח הודעה יחסית ארוכה ומקבל תגובה יחסית קצרה מאחד חברי הקבוצה כחמש שניות לאחר מכן, התוקף כמובן עונה לו עם הודעה קצת יותר קצרה מהראשונה ששלח אבל עדיין יחסית ארוכה כ 15 שניות לאחר תחילת ההקלטה. השיחה ממשיכה בכך שחבר בקבוצה שולח הודעת טקסט לאחר 27 שניות ומקבל תגובה ארוכה מאוד מהתוקף (ניתן לראות כי ההודעה נשלחה במספר פקטות נפרדות כיוון שגודל הפקטה המקסימלית היא 1500) לאחר מכן חבר הקבוצה שלח הודעה קצרה (סביב ה 3 שניות לאחר מכן) והתוקף שלח גם הוא הודעה קצרה שלה ענה חבר הקבוצה בהודעה ארוכה כ 60 שניות לאחר תחילת ההקלטה.

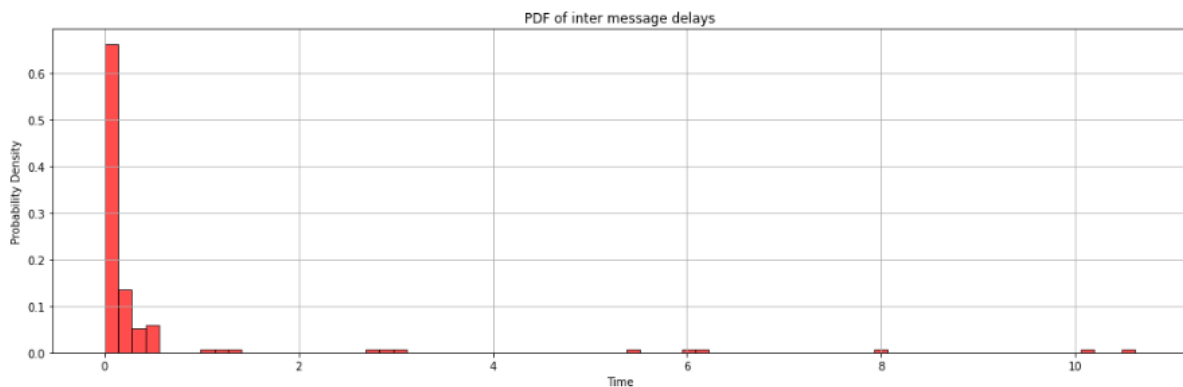
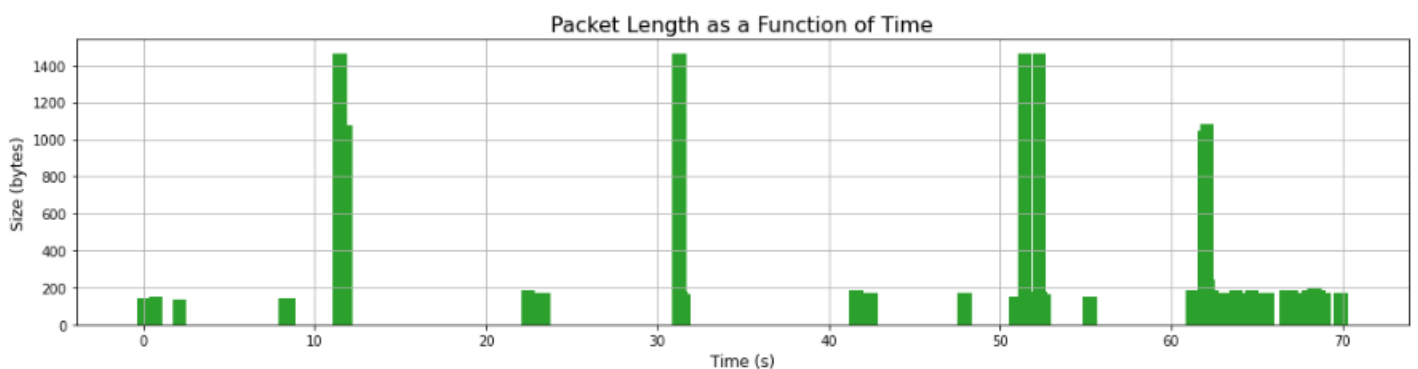
נשים לב שיש המון רעש בגרף וזה נובע מהעבודה שההקלטה הוקלטה בשעת אחר הצהריים ולכן התקבלו עוד הודעות מקבוצות בהן התוקף חבר.



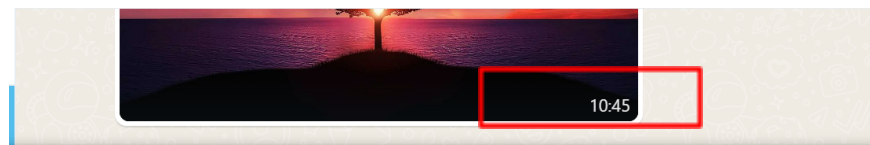
הקלטה 2: תמונות

בקבוצה השניה נשלחות רק תמונות ניתן לראות שהתוקף מקבל תמונה ראשונה עשר שניות אחרי שמתחיל את ההקלטה ושולח תמונה חזרה 12 שניות לאחר מכן. התוקף שוב מקבל עוד תמונה סביב ה 30 שניות ומחזיר תמונה עשר שניות אחרי. חבר הקבוצה שולח שתי תמונות רצופות שעליהן עונה התוקף עם מספר רב של תמונות.

נשים לב כי התמונות שהתוקף שולח מתחלקות להמון פקטות קטנות על מנת להעביר כמה שיותר מידע כמה שיותר מהר אבל כשהתוקף מקבל תמונות הוא מקבל את הפקטות בגודל המקסימלי האפשרי ולכן ניתן לראות שהחבילות הרבה הרבה יותר גדולות.



נרצה לוודא כי אנחנו באמת מקבלים את ההודעות הנכונות בזמן הנכון לכן נצרף את הצילום מסך הבא:



es shark · Packet 123 · Wi-Fi

Interface id: 0 (\Device\NPF_{D83FE77D-48E8-405B-940D-BD5533BD083F})

Encapsulation type: Ethernet (1)

Arrival Time: Aug 15, 2023 10:45:10.357540000 Jerusalem Daylight Time

[Time shift for this packet: 0.000000000 seconds]

Epoch Time: 1692085510.357540000 seconds

[Time delta from previous captured frame: 0.000000000 seconds]

[Time delta from previous displayed frame: 0.000000000 seconds]

[Time since reference or first frame: 0.935701000 seconds]

הקלטה 3: הודעות קוליות

בהקלטה הנוכחית התוקף נמצא בקבוצה בה שולחים רק הודעות קוליות, בניסוי שלנו אנחנו מניחים כי אנשים שולחים הודעות קוליות של לפחות 5 שניות וזה מהסיבה שהודעות קוליות לרוב משמשות כאשר מישהו רוצה להגיד משהו שארוך מידי כדי לכתוב לכן פונה לאופציה הנוחה יותר.

בגרף ניתן לראות כי התוקף מתחיל את השיחה כ 15 שניות לאחר שהתחיל את ההקלטה ושולח הודעה קולית יחסית ארוכה שנפרסת על כמה פקטות (מהסיבה שפקטה מקסימלית הינה באורך של 1500 בייטים), לאחר ששולח הודעה התוקף מקבל תשובה כ 15 שניות לאחר מכן, שוב ניתן לשים לב שמקבל תשובה יחסית ארוכה שלה עונה בהודעה קולית קצרה יותר במעט. אנחנו ממשיכים לראות עוד פיקים כל 15 שניות עד סיום ההקלטה.

בהקלטה זאת הוספנו את ההקלטות שביצענו משני המחשבים (של שני הצדדים) וזה כיוון שעשינו ניסוי סביב 75 שניות לאחר תחילת ההקלטה. כפי שניתן לראות יש פיק רק בגרף הראשון בעוד שבשני אין שכן לעובדה שהתקבלה איזושהי הודעה וזה קרה כיוון שהתוקף הקליט הודעה אך מחק אותה במקום לשלוח אותה. עובדה זאת מעניינת כיוון שבעזרת טכניקה זאת התוקף לא צריך לחשוף את עצמו אלא רק לראות מתי יש אצלו פקטות שהוקלטו מבלי לשלוח הודעות שיכולות להיות כביכול רנדומלית בקבוצה.



נקודה חשובה

במהלך העבודה שמנו לב שאם אנחנו מקליטים הודעה קולית או לחלופין כותבים הודעת טקסט יש הקלטה של פקטות יוצאות מהמחשב של השולח אבל אין זכר לפקטות נכנסות בצד השני.

תמונות

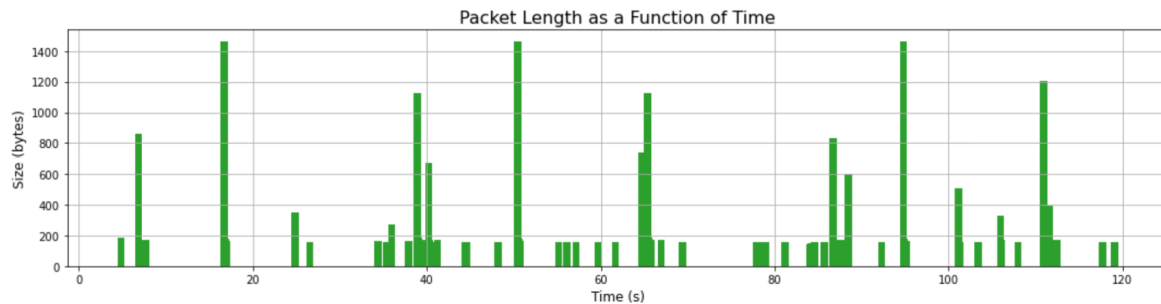
Maya Univ is typing...

אנחנו מסיקים מכך שכן עוברות פקטות כלשהן כי הרי הצד השני כן יודע שאחד חברי הקבוצה כותב ברגע זה אבל ההודעה עצמה לא עוברת למרות שהיא מוכנה לצאת.

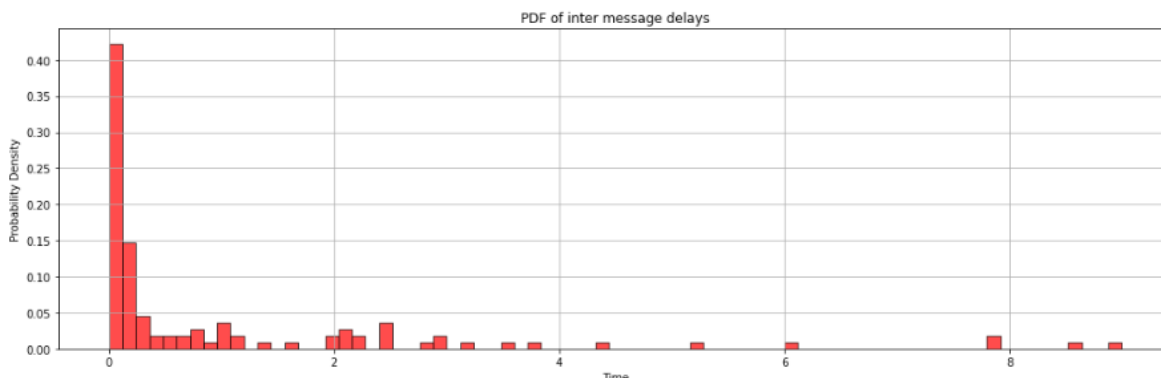
מנגד, שמנו לב שדבר זה לא קורה במידה ואנחנו מעלים תמונה לשיחה אבל מוחקים אותה מבלי לשלוח, אין שום הודעות שמתקבלות אצל הצד השני.

הקלטה 4: סרטונים

בקבוצה הרביעית שלחנו סרטונים בלבד על מנת להבין כיצד שליחת סרטונים בקבוצה משפיעה על יכולת ההבנה של התוקף לגבי מה שקורה בקבוצה. כפי שניתן להבין מהגרף המצורף הסרטון הראשון התקבל כ-15 שניות לאחר תחילת ההקלטה והתוקף שלח סרטון יחסית ארוך לאחר 25 שניות שהתחלק למספר פקטות. לאחר כ-10 שניות התקבל עוד סרטון מחבר הקבוצה (בדיוק אותו סרטון כפי שנשלח בפעם הראשונה) והתוקף שלח סרטון שונה לאחר 15 שניות. התוקף שלח עוד סרטון דקה וחצי אחרי תחילת ההקלטה וחבר הקבוצה שלח בפעם השלישית את אותו הסרטון. רגע לפני שסיים התוקף את ההקלטה הוא החליט לשלוח סרטון אחרון.



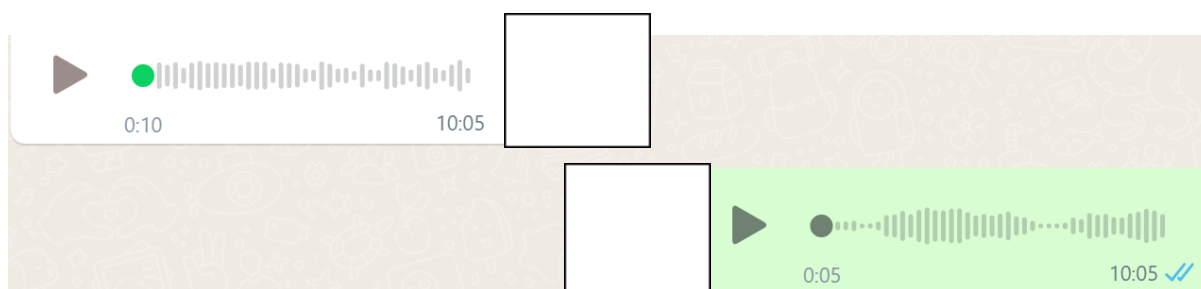
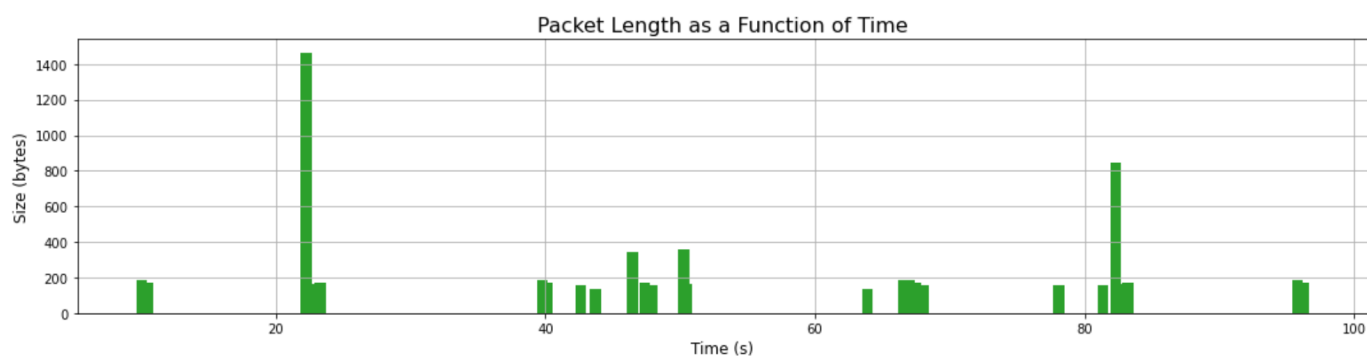
החלק של הסרטונים היה המאתגר ביותר כיוון שכל פעם כל סרטון מתפצל לכמות שונה של פקטות ולפעמים לוקח לסרטון אפילו כמה שניות לטעון בשונה מתמונות, הודעות טקסט והודעות קוליות שלוקח מקסימום שניה אחת להודעות לטעון.



נשים לב בהתאם למה שכתוב לעיל הגרף של הדיליי אכן שונה מה-PDF של הקבוצות הקודמות כלומר יש דיילים ארוכים יותר

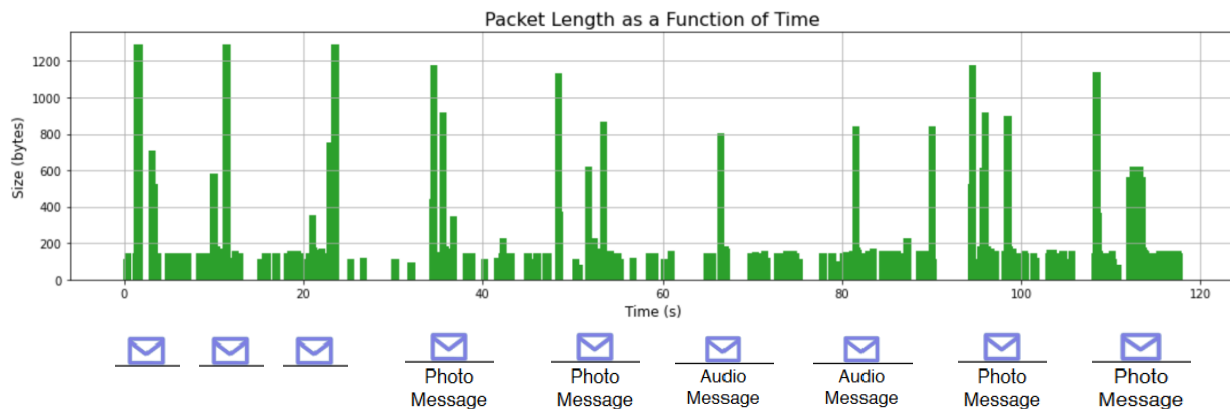
הקלטה 5: משלבים הכל ביחד

התוקף מתחיל שיחה בקבוצה כ 10 שניות לאחר שהפעיל את ההקלטה ושולח הודעה קצרה, לאחר שמקבל תשובה (תמונה - כפי שניתן לראות לפי העובדה שהפיק גבוהה יותר ב 20 שניות) הוא עונה עם הודעת טקסט ארוכה יותר ב 40 שניות ומקבל תשובה (בטקסט) ב 50 שניות. לאחר מכן התוקף שולח 2 תמונות ברצף לקבוצה (סביבות ה 65 שניות) ומקבל הודעה קוליות כ 15 שניות לאחר מכן מהצד השני ושולח בחזרה הודעה קוליות קצרה בהרבה (5 שניות של הקלטה לעומת 10 שניות) ומסיים את השיחה ואת ההקלטה.



הקלטה 6: הוספת רעש להקלטות

בהקלטה השישית והאחרונה הוספנו רעש להקלטה או יותר נכון לא פילטרנו את ההקלטה בכלל. תחילה פתחנו Discord ודיברנו דרך האתר ובנוסף פתחנו YouTube. נכנסנו לקבוצה ושלחנו הודעות מעורבות, תחילה שלחנו הודעות רגילות בקבוצה ולאחר מכן (סביב ה 22 שניות) התוקף שלח תמונה ולאחר 10 שניות קיבל תמונה חזרה מחבר הקבוצה. לאחר מכן שלח התוקף הודעה קולית ארוכה יחסית וקיבל חזרה הודעה קולית (סביבות ה 65 שניות) קצרה יותר. בסוף התוקף שלח שתי תמונות וקיבל מחבר הקבוצה 6 תמונות.



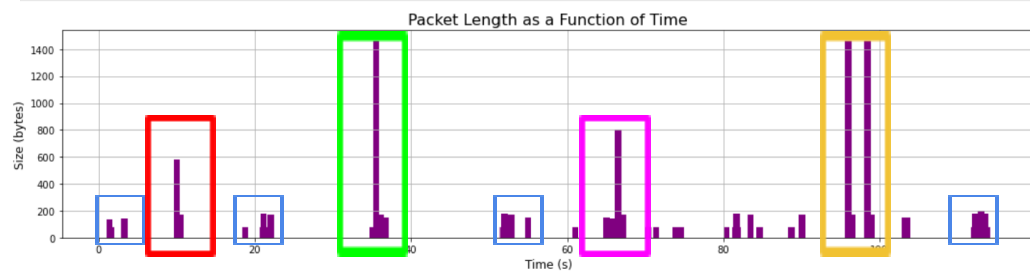
ניתוח התעבורה אצל שני הצדדים

על מנת לראות רק את ההודעות הנשלחות או רק את ההודעות הנכנסות נצטרך להשתמש בפילטר הבא Wireshark.

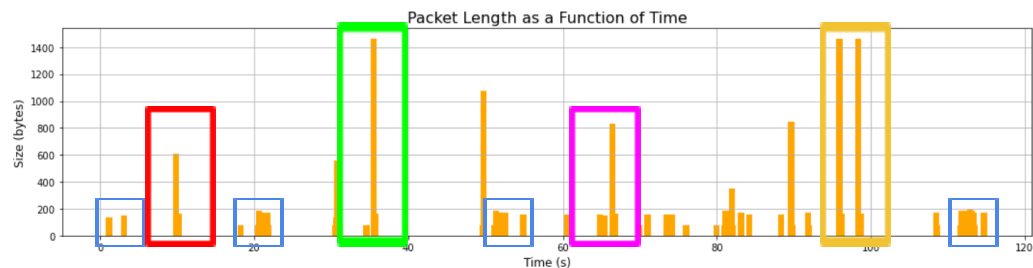
```
tcp.port == 443 && ipv6.dst == 2a03:2880:f258:c8:face:b00c:0:167
```

```
tcp.port == 443 && ipv6.src == 2a03:2880:f258:c8:face:b00c:0:167
```

```
In [17]: displaySizeVsTime('AttackerSent.csv','purple')
```



```
In [18]: displaySizeVsTime('attackedRecv.csv','orange')
```



בחלק האחרון של העבודה רצינו להראות בדיוק איך עובד תעבורת הפקטות, בגרף מעל אנחנו רואים בדיוק איך הפקטות שנשלחות מהתוקף הן חלק מהפקטות המתקבלות אצל חבר הקבוצה (כמובן שיש עוד פקטות שקיבל כיוון שזאת לא הקבוצה היחיד שהוא חבר בה לפי ההנחה בעבודה). ההקלטות הוקלטו משני המחשבים של התוקף וחבר הקבוצה והתחילו בדיוק באותה שניה על מנת שיהיה ניתן להבחין במדויק בהודעות הנשלחות והמתקבלות אצל שני הצדדים.

נשים לב שישנו רעש שחוזר על עצמו בשני הגרפים וזה מהסיבה שפה לא סיננו את ה ACKים לכן אנחנו רואים בהקלטה את כל לחיצות הידיים שקורות במהלך שליחת ההודעות.

סיכום

בפרויקט חקרנו את התקפות על קבוצות של הודעות באפליקציות שונות לשליחת הודעות מאובטחות, תחילה קראנו את המאמר ולקחנו איתנו את הנקודות החשובות ומימשנו בפרויקט, כלומר בחלק הרטוב של העבודה. תחילה פתחנו 4 קבוצות שונות ב-WhatsApp Web והתחלנו את ההקלטות דרך ויירשארק. בכל פעם רשמנו לעצמנו מי שלח איזה סוג של הודעה ומתי (השתמשנו בסטופר) וכך ידענו לנתר בקלות את ההודעות בהקלטות ויכולנו להצביע בדיוק על איזה הודעות נשלחו בקבוצה (בכוונה כחלק מהניסוי) ואיזה הודעות התקבלו בקבוצות אחרות ומהוות רעש בהקלטה הנוכחית.

התוצאות שלנו הראו כי התקפות ניתוח תנועה יכולות לשמש לזיהוי משתמשים ופעילויות ב-WhatsApp, גם אם ההודעות עצמן מוצפנות. לדוגמה, לפי הנתונים שהקבלו ובהנחה שהתוקף כבר חבר בקבוצות ההודעות (בין שהקבוצה הייתה פתוחה/ציבורית והוא יכול היה להוסיף את עצמו, או שמישהו שכבר היה מנהל בקבוצה ספציפית הוסיף אותו), הצלחנו לזהות אילו משתמשים שולחים תמונות, הקלטות קוליות והודעות טקסט ובדיוק באיזה זמן (אנחנו יודעים בדיוק באיזה שניה התחילה ההקלטה וניתן לראות במדויק כפי שראינו בהתחלה באיזה שניה התקבלה כל פקטה). התוצאות שלנו מראות שהתקפות ניתוח תנועה הן איום רציני על אבטחת אפליקציות לשליחת הודעות מאובטחות.