פרויקט גמר רשתות תקשורת

<u>מגישות:</u>

אביה אורן 322273301 נטע כהן

: קישורים ללינקדאין

www.linkedin.com/in/aviya-oren-80471926b : אביה

: נטע

<u>קישור לgithub של הפרויקט:</u>

https://github.com/OrenAviya/Networks_FinalProject

חלק יבש:

Practical Traffic Analysis Attacks on Secure" סיכום המאמר "Messaging Applications

:הרעיון המרכזי המובא במאמר הוא

המאמר מתאר דרך למעקב והשגת מידע רגיש מתוך התכתבויות באפליקציות הודעות, במאמר מתואר חקר על יכולת הבקרה והמעקב על אנשים דרך ההודעות באפליקציות הודעות IM = instant מתואר חקר על יכולת הבקרה והמעקב על אנשים דרך ההודעות באפליקציות למשל whatsapp או telegram . למרות ההצפנה שבד"כ משתמשים בה באפליקציות כגון אלו ניתן לבצע מעקב וניתוח התעבורה כך שמתאפשר זיהוי מנהל / משתתף בקבוצה מסויימת יריבה וזיהוי פעילים ומשתתפים בה.

לרוב שירותי הודעות מיידיות ראשיים אינם מטשטשים דפוסי תעבורה כי זה יקר לביצוע. ודרך דפוסים אלו אפשר לגלות על השולח כל מיני דברים.

כותבי המאמר מקווים שהתוצאות שהשיגו במחקר שלהם ובפיתוח שרת הפרוקסי שיכול להגן על מעבירי ההודעות יזהירו את ספקי הודעות מיידיות ויגרמו להם לקחת אמצעים יעילים יותר להגן על לקוחותיהם.

ראשית נתייחס לשאלה - כיצד התוקף משיג נתונים בסיסיים "ground truth" על התעבורה של הערוץ? וכיצד הוא מדייק את הנתונים שלו מול המציאות?

- מקרה ראשון אם התוקף יכול להצטרף לערוץ אליו הוא רוצה להאזין (איפה שנשלחות ההודעות האטרקטיביות לבדיקה מבחינתו) ולצפות במה שמעבירים באותו ערוץ הוא יכול להשוות את הנתונים שהוא רואה לנתונים שהוא משיג ולגלות מה הדיוק של הנתונים שהוא לוכד.
 - 2. מקרה שני התוקף יכול להצטרף כמנהל ואף ל**שלוח הודעות** בעצמו בקבוצה בנוסף על היותו **צופה.**
- 3. במקרה הגרוע- התוקף לא יכול ממש להצטרף לערוץ אבל יכול לזהות משתתפים ומנהלים דרך תקשורת אחד על אחד . כלומר אם יצליח לזהות כתובת פרטית של אחד ממשתתפי הקבוצה ולנתח את התעבורה שלו , יוכל בהנחה שהוא יודע שזהו אחד מהחברים בקבוצה, להשליך מתוך ניתוח התעבורה הפרטית שלו על המסרים שנשלחים בערוץ- בקבוצה.

כאמור לעיל הניתוח כולו מתבצע באמצעות הגודל והתזמון/תדירות של החבילות שעוברות ברשת , היות והמסרים עצמם מוצפנים.

כיצד מתבצע הניתוח בפועל?

באופן כללי - הגודל והתזמון של חבילות יכולות להדליף מידע על זהות השולח או המקבל בגלל הדמיון בין החבילות שנשלחו מהמשתמש שנבדק והתקבלו בערוץ (אם זו אותה סוג של הודעה יש לה גודל כמעט שווה והזמנים יהיו דומים אזי ניתן להניח שזו השתתפות פעילה של האדם השולח בקבוצה הנבדקת.)

באופן יותר ספציפי- אפשר לבצע מעקב זה תוך שימוש בשני אלגוריתמים אפשריים-

Event-Based Algorithm.1

כלומר, אם שני "אירועים" / הודעות אחד בתעבורת הערוץ ואחד בתעבורה שנשלחה מאדם מסויים נשלחות מספיק "קרוב" בזמן הוא יוצר התאמה ואם יש התאמות מרובות ביחס לכמות ההודעות שנשלחו בכלל (בהשוואה לrashhold שנקבע מראש) אזי התוקף יודע ל"סמן" את האדם כפעיל באותו ערוץ תקשורת.

shape-Based Algorithm.2

אלגוריתם זה מבוסס על "צורת" התעבורה וה"צורה" היא בעצם וקטור של אורכי מנות לאורך זמן. כלומר לפי הצורה שיוצרות החבילות - מתאפשר לו להשוות ולזהות חבילות מאותו סוג שנשלחו והגיעו לכל המשתתפים בערוץ.

איך מחושבת הצורה - הווקטור הנ"ל?

זיהוי " אירוע" - הודעה מסויימת ואז נרמול התעבורה לפי חבילות בגובה שווה. הצורה הסופית מתקבלת מהווקטור שנוצר ממעבר על גבהי האירועים לאורך זמן ההקלטה. גם כאן - ההשוואה היא בין התעבורה של משתמש פרטי למול התעבורה הכללית בקבוצה.

נתייחס לשאלה נוספת - כיצד התוקף מאזין בסתר לתעבורת הרשת? כיצד הוא מגלה את כתובת הריחס לשאלה נוספת - כיצד התוקף מאזין לבדוק?

ישנן כמה דרכים לבצע זאת:

- 1. האזנת סתר לתעבורת הרשת של ספקי האינטרנט או ה-IXP שבהם התוקף שולט, למשל, אם זהו ארגון ממשלתי ששולט בחלק מהרשת. דוגמא לכך יכולה להיות חומת האש הגדולה של סין.
 - 2. לחלופין, היריב יכול להאזין לרשת תעבורה של אנשים ספציפיים (למשל, פעילים חשודים), לאחר קבלת צו האזנת סתר.

כותבי המאמר רצו לדמות "התקפה" של יריב (למשל ממשלה) על תקשורת רגישה כלשהי (למשל פוליטית) . ולכן מידלו את התעבורה בערוצי רשת ומיינו אותה לפי חמשת סוגי ההודעות הנמצאים בשכיחות הגבוהה ביותר בתעבורה ברשת : קבצים, תמונות , סרטים, אודיו (שמע) , והודעות טקסט.

הנתונים סוכמו לטבלה 2 במאמר:

TABLE II: Distribution of various message types

Type	Count	Volume (MB)	Size range	Avg. size
Text	12539 (29.4%)	3.85 (0.016%)	1B-4095B	306.61B
Photo	20471 (48%)	1869.57 (0.765%)	2.40Kb-378.68Kb	91.33KB
Video	6564 (15.4%)	232955.19 (95.3%)	10.16Kb-1.56Gb	35.49MB
File	903 (2.1%)	47.46 (0.019%)	2.54Kb-1.88Mg	52.56KB
Audio	2161 (5.1%)	9587.36 (3.92%)	2.83Kb-98.07Mg	4.44MB

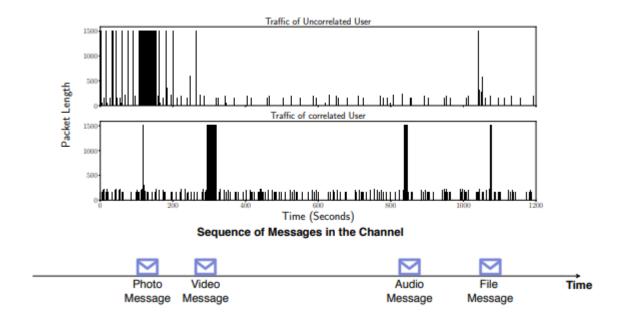
בטבלה נבחנו ההודעות לפי 4 פרמטרים:

- 1. כמות ההודעות מהסוג הזה.
- 2. שטח האחסון שסוג זה תפס בזיכרון
- 3. טווח הגדלים שהודעות מסוג זה יכולות לקבל
 - 4. גודל ממוצע

: סיכום הטבלה

- הודעות טקסט ותמונות תופסים אחוז נכבד מסך כל התעבורה. (תמונות עם 48% והודעות טקסט עם 29.4%) ואחריהם מגיעים וידאו (15.4%), שמע (5.1%) וקבצים (2.1%)
 - . (וידאו) ע"י סרטונים וסרטים (וידאו) נפח האחסון נתפס רובו
- טווח הגדלים של תמונות והודעות טקסט קטן יחסית לטווח הגדלים של שמע וקבצים וכל אלו
 קטנים מאוד לעומת טווח הגדלים שוידאו יכול לקבל הנע בין 10KB ו1.56GB!
- גודלם הממוצע מהקטן לגדול של מסרים מכל סוג הוא: הודעת טקסט: 306.61B , קובץ (קובץ 35.49MB , שמע: 4.44MB , שמע: 91.33KB

לאחר מידול זה של התעבורה הכללית ברשת, כותבי המאמר חקרו את היכולת לחלץ מידע על משתתפים בערוץ תקשורת מסוים.



איור 8 במאמר , מדגים איך ניתן לזהות קורלציה בין תעבורה של משתמש שלא שייך לקבוצה מסויימת (הגרף העליון) לבין תעבורה של משתמש ששייך לקבוצה מסויימת (הגרף התחתון) ע"י השוואה של התעבורה שלהם לתעבורת הערוץ.

?איך זה מתבצע

עבור תעבורה בערוץ מסוים ניתן לראות את המרחקים בין החבילות שנשלחו וכן את גודל החבילות וכך לזהות אילו חבילות שייכות לאותה ההודעה/ "אירוע". זהו שלב זיהוי "אירוע מסויים" - רצף פאקטות שמייצגות מסר אחד שנשלח בקבוצה.

התוקף יכול לחלץ מהתקשורת המוצפנת צרורות של פאקטות אלו , לבדוק האם הן נשלחו באותו "פרק זמן" (טראשהולד שנקבע מראש) ולקבוע שזהו "אירוע" אחד, כלומר שהתעבורה הזו שייכת להודעה אחת.

לאחר מכן יוכל להשוות אירועים אלו לאירועים מקבילים שהתרחשו באותו זמן בתעבורה שנלכדה אצל משתמש ספציפי ולקבוע האם הוא ממשתתף בערוץ או לא.

הדרך בה מנסים כותבי המאמר להתמודד עם מעקב שכזה היא יצירת שרת פרוקסי דרכו מועברות חבילות התעבורה שנשלחות מהמנהל ומתקבלות אצל המשתתפים. מה שהפרוקסי עושה זה "מחיקת" הדפוס שיצר השרת המקומי של השולח ושינוי ה"זמנים" או הוספת פקטות מדומות כך שלא יוכלו העוקבים לקשר בין החבילות שנשלחו לכדי "אירוע אחד".

הם גם מוכיחים שבאמצעות אלגוריתם זה הם מצליחים להפחית את יכוולת המעקב של "תוקף" ביותר מ20%.

חלק רטוב- ניתוח תעבורה בקבוצות:

: נדרש לבצע

1. להראות הבדלים בקבוצות שונות שנשלחים בהם בד"כ סוגי הודעות שונים (גדולים יותר ופחות למשל) (אפשר להניח שיש לנו מרגל בקבוצה שיודע אילו קבצים הועברו ומתי)

:הקדמה למחקר

בחרנו להאזין לקבוצות בווצאפ. (מתוך הנחה שאנחנו נמצאות בקבוצות).

ע"מ לדעת איזה סינון ידרש ממנו כדי לזהות את הפאקטות הנכונות חיפשנו מה כתובת האמנו לדעת איזה סינון ידרש ממנו כדי לזהות את הפאקטות הנכונות חיפשנו מה כתובת whatsappweb:

C:\Users\97258>nslookup web.whatsapp.com

Server: UnKnown

Address: fe80::a8ab:b5ff:fe32:5a64

Non-authoritative answer:

Name: mmx-ds.cdn.whatsapp.net

Addresses: 2a03:2880:f242:c8:face:b00c:0:167

157.240.195.56

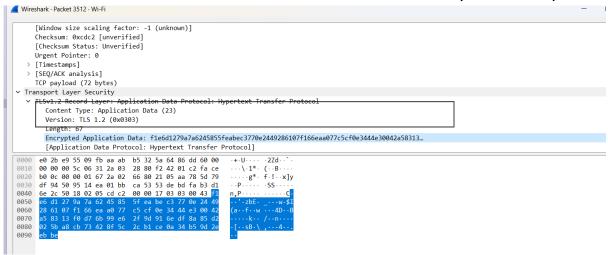
Aliases: web.whatsapp.com

את הכתובת הזו אנחנו רואות בכל פעם שמופיעה הודעה מווצאפ: בתמונה ניתן לראות את

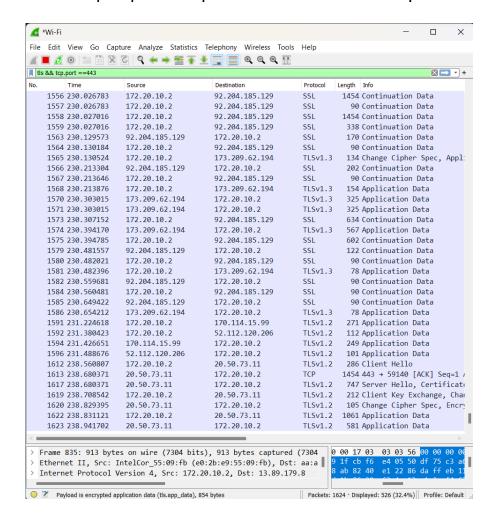
הסינון שבחרנו כדי לתמצת כמה שיותר את התעבורה לכדי הודעות בלבד (ועדיין ישנם נתמודד איתם). אנחנו בוחנות רק הודעות המתקבלות בווצאפ. ניתן לראות שהשרת זהה עבור כל ההודעות על אף שהן מתקבלות

4	*Wi-Fi																-			×
File	Edit	View	Go	Captu	ire /	Analyze	Stati	stics	Teleph	ony	Wirel	less	Tools	Help						
		②		X	5	₹ 🛻 🗷	**	* ₹			⊕ €		* *							
t	ls.record.	version	== 0x0	303 &&	tcp.p	ort==443	8& ip	/6.src =	= 2a03:	2880:f	242:10	c2:fac	e:b00c:0	:167					×I→	-
No.		Time			Sour	ce			Des	tination				Protocol	Length	Info				
	3145	522.	20295	9	2a03	3:2880	:f242	:1c2:	2a0	2:66	80:2	105:	aa78	TLSv1.3	245	Appl	ication	Data		
	3146	522.	20295	9	2a03	3:2880	:f242	:1c2:	2a0	2:66	80:2	105:	aa78	TLSv1.3	148	Appl	ication	Data		
	3151	522.	21539	8	2a03	3:2880	:f242	:1c2:	2a0	2:66	80:2	105:	aa78	TLSv1.3	131	Appl	ication	Data		
	3152	522.	25951	0	2a03	3:2880	:f242	:1c2:	2a0	2:66	80:2	105:	aa78	TLSv1.3	565	Appl	ication	Data		
	3155	522.	56556	0	2a03	3:2880	:f242	:1c2:	2a€	2:66	80:2	105:	aa78	TLSv1.3	312	Appl	ication	Data		
	3207	537.	95932	7	2a03	3:2880	:f242	:1c2:	2a0	2:66	80:2	105:	aa78	TLSv1.2			ication			
			69230											TLSv1.2			ication			
			30521	-										TLSv1.2			ication			
			87771											TLSv1.2			ication			
			06332	-										TLSv1.2			ication			
			69506	-										TLSv1.2			ication			
	3512	610.	52573											TLSv1.2			ication			
			59254											TLSv1.2			ication			
			69069											TLSv1.2			ication		Apr	
			56959											TLSv1.2			ication			
			84228											TLSv1.2			ication			
			76599											TLSv1.2			ication			
			74057											TLSv1.2			ication			
			00148	-			– . –							TLSv1.2			ication			
			65047											TLSv1.2			ication			
			65047											TLSv1.2			ication			
			65047											TLSv1.2			ication			
			73522	-										TLSv1.2			ication			
			74216	-										TLSv1.2			ication			
			74216	-			– . –							TLSv1.2			ication			
			75429	-			– . –							TLSv1.2			ication			
			81630											TLSv1.2			ication			
			31471				– . –							TLSv1.2			ication			
			61726											TLSv1.2			ication			
	3724	642.	72853	5	2a03	3:2880	:f242	:1c2:	2a0	2:66	80:2	105:	aa78	TLSv1.2	134	Appl	ication	Data		

דבר נוסף וחשוב להבין - הודעות ווצאפ מוצפנות כמו בדוגמא:



ע"מ להגיע לסינון הזה , הקלטנו הקלטה של פעילות המחשב כאשר האפליקציה ווצאפ סגורה, וסיננו את התעבורה לפי מסנן tls והפורט הנכון וכך הגענו למסקנה שהתעבורה משרתים בצורת ipv4 שאנו רואות אינה קשורה לווצאפ וכן שרתים נוספים שאינן השרת של ווצאפ ניתן לראות בדוגמא את צילום המסך ווירשארק במקרה כזה:



נסיון ראשוני:

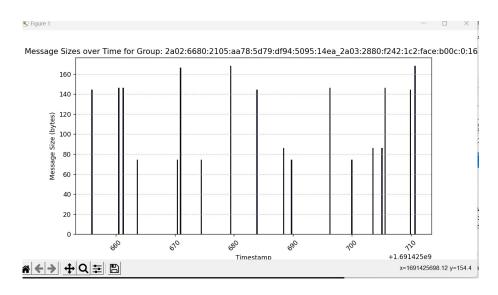
נשאלת השאלה- איך נדע להבחין ולהסיק מסקנות מהמידע המצומצם שיש לנו בווירשארק (גודל הפאקטה וזמן השליחה)?

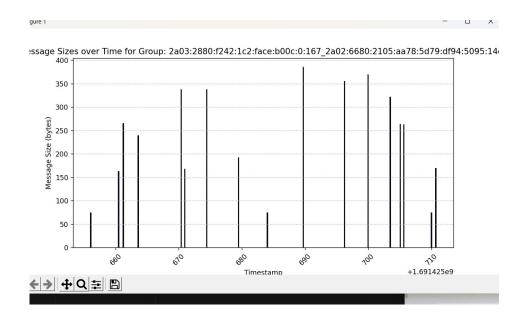
ניסינו לתפוס את ההבדלים בתעבורה בין סרטונים תמונות והודעות טקסט קצרות. פתחנו קבוצה והקלטנו בווירשארק את התעבורה מהשרת של ווצאפ (שגילינו לפי סינון פרוטוקול tcp , הפורט המתאים והשרת שמצאנו) וזיהוי הודעות שנשלחו מאפליקציה כאשר זו הייתה האפליקציה היחידה במחשב שלנו.

בנוסף כתבנו קוד שיקלוט את התעבורה בפרוטוקול TCP במשך דקה, יחלק אותה לקבוצות לפי המקור והיעד וגם ויציג לנו אך ורק את קבוצות ההודעות שמקורן במחשב שלנו ויעדן בשרת של וואטסאפ או להפך.

כך יצרנו גרפים המראים את הפאקטות שעוברות והפרטים המוצגים לגביהן הם הזמן שעבר בין פאקטה לפאקטה והגודל שלה (בדומה לגרפים שהוצגו במאמר.)

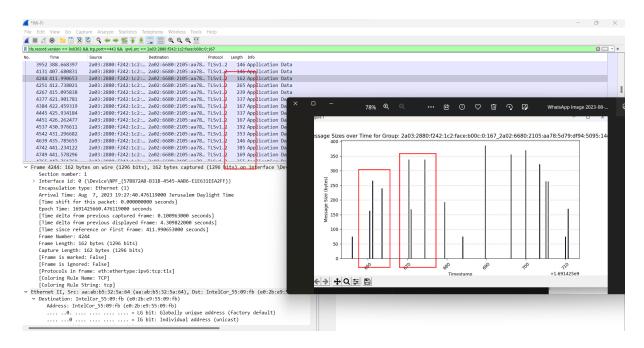
למשל הגרפים לאחר דקת תקשורת קצרה בהודעות טקסט בקבוצה שלנו:





ניסינו להתאים בין מה שראינו בווירשארק לבין הגרף שהציג הקוד.

בצילום המסך שלמטה (איור 6) רואים קשר בין גודל הפאקטות בווירשארק וכמובן המקור והיעד לבין הגרף (ישנן גם פאקטות שלא חופפות)

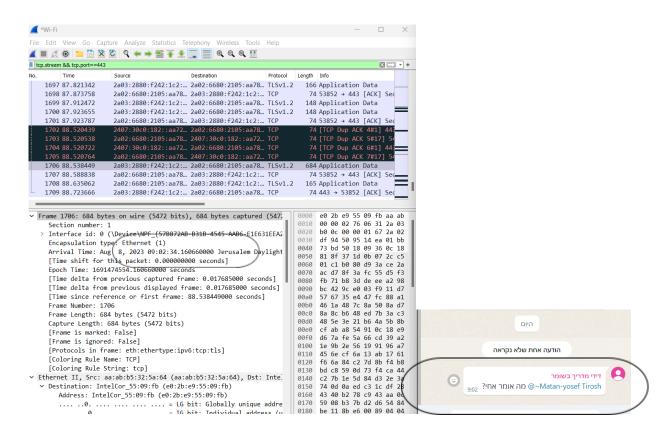


(איור 6): נביט מפאקטה בגודל 162 שהיא העמודה השניה בגרף שלנו לאחריה 239 265 וגם 337 מתאימות וכן אפשר לבדוק הלאה. ישנן פאקטות שמופיעות בווירשארק אך לא נכנסו לגרף. ההשערה שלנו היא שהסינון של הקוד לא מדוייק ונכנסות בו גם פאקטות שהן "רעשי רקע" מבחינתנו ולכן ההתאמה לא מדוייקת.

מצאנו התאמה בין גודל הפאקטות אך עדיין חסר לנו מידע לגבי איך הן מתחלקות ותוך כמה זמן נשלחת הודעה שלמה

בהמשך למחקר שלנו - ניסינו לעקוב אחד "stream" אחד כדי לדעת אילו הודעות שייכות לאותה שיחה

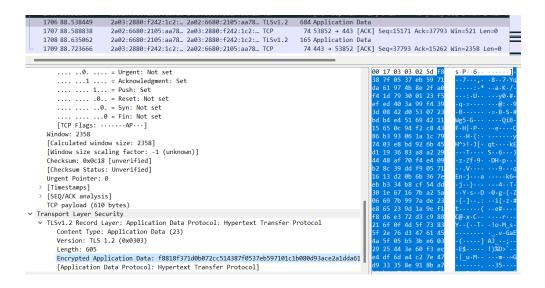
ור אותה ע"מ לדעת לסנן אך ורק הודעות שאינן לחיצות;



רואים שהזמנים חופפים בין ההודעה והפאקטה שנתפסה ורואים שלפניה ואחריה זוהי כבר לא אותה הודעה כי הפאקטות לפניה ואחריה מקורן לא בשרת ווצאפ שממנו מגיעות ההודעות אלא בשרת של המחשב שלנו מול ווצאפ.

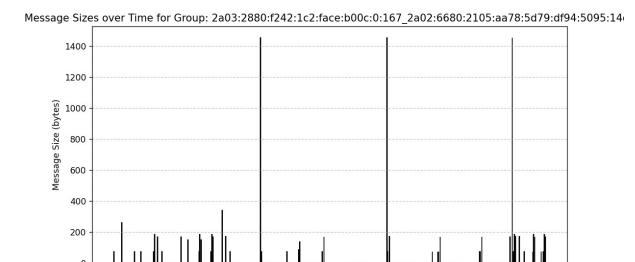
נשים לב למאפיינים - הפרוטוקול הוא TLS גרסה 1.2 בהוא אחראי להצפנת הData. בנוסף הפורט והשרת תואמים לגילויים הקודמים שלנו לגבי שרת ווצאפ- המשתמש בפורט ipv6: Src: 2a03:2880:f242:1c2:face:b00c:0:167 Dst: 2a02:6680:2105:aa78:5d79:df94:5095:14ea

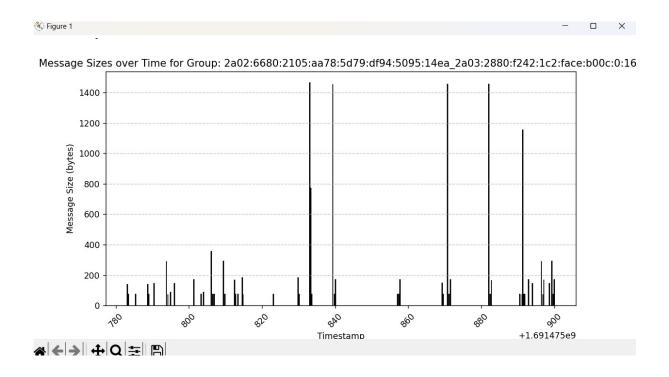
בתיאור כתוב לנו "application data" ובנוסף אפשר לראות בתוכן הפאקטה את הדאטה המוצפן:



אז בדוגמא זו ההודעה הייתה כנראה מספיק קצרה כך שהיא נשלחה בחתיכה אחת. באורך 605 בייטים.

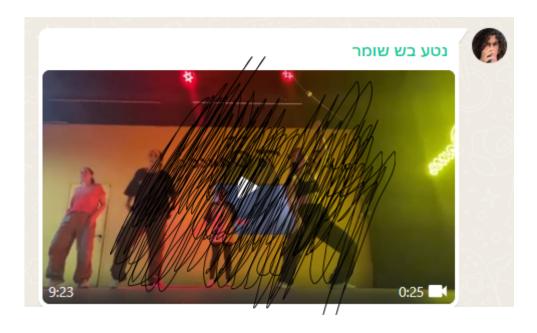
על מנת לבדוק כמה פאקטות שייכות לאותה הודעה ואיך נראה ההבדל בין סרטונים ותמונות להודעות רגילות פתחנו שיחה בקבוצה שלנו (בזמן שלא היו הודעות אחרות באפליקציה) ואלו התוצאות אחרי 2 דקות שנתנו לקוד לרוץ ובתוכן שלחנו תמונות וסרטונים בקבוצה:





בהקלטת הווירשארק המתאימה ראינו כמה חפיפות:

למשל בשעה 9:23 לקראת סופה נצפו 3 פאקטות גדולות המקבילות בערך לגודל 1450 ובדיוק בזמן זה נשלח סרטון בקבוצה שבחנו:



סיכום ביניים- ממצאים שראינו עד עכשיו:

שרת ווצאפ במחשב שלנו - שממנו ואליו מגיעה התעבורה אותה אנחנו בוחנות הוא:
 2a03:2880:f242:1c2:face:b00c:0:167

443 : הפורט המתאים

- ניתן לראות שהשרת זהה עבור כל ההודעות על אף שהן מתקבלות מקבוצות שונות.
- ההודעות המתקבלות משרת זה ואליו מוצפנות ע"י פרוטוקול TLS כלומר אי אפשר לדעת מה תוכן ההודעה רק את גודלה ואת הזמן שנשלחה.
 - הודעות טקסט לעיתים נשלחות בפאקטות גדולות וסרטונים ותמונות ככל הנראה מחולקים לכמה פאקטות גדולות.

: לאחר נסיונות אלו שיפרנו את הקוד

הקוד גם ישמור קובץ טקסט ובו כל הפאקטות שתפס.

המחקר עצמו - הקלטת 4 קבוצות:

לצורך המחקר, לפי מה שהתבקשנו, פתחנו 4 קבוצות שונות , בכל קבוצה נשלח תוכן אחר שיהיה העיקרי .

נקליט 10 דקות כל קבוצה ונבדוק את ההבדלים בין הגרפים שיוצגו.

בנוסף נבדוק ונשווה את הקלטות הווירשארק ואת המידע שתפס הקוד שלנו. ונבדוק סטטיסטיקות של הפרשי זמן בין הודעות. בעזרת קובץ EXCEL אליו נייצא את תוצאות הווירשארק

קבוצה 1 - הודעות טקסט

קבוצה 2- תמונות

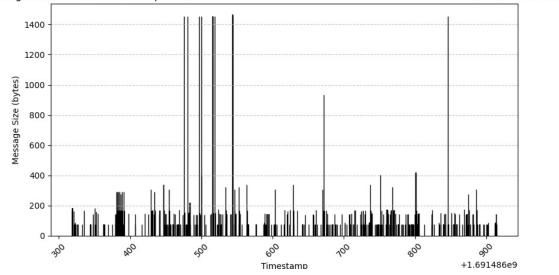
קבוצה 3 - סרטונים

קבוצה 4 - שמע, הקלטות קוליות.

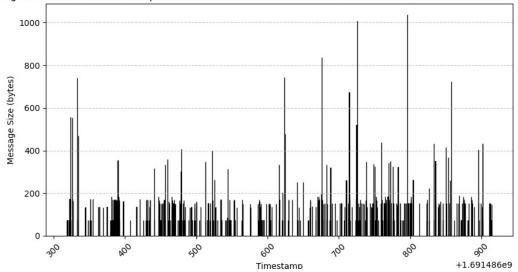
:האזנה מספר 1 - הודעות טקסט

10 דקות של הודעות טקסט בקבוצה תוצאות:

Message Sizes over Time for Group: 2a02:6680:2105:aa78:5d79:df94:5095:14ea--2a03:2880:f242:1c2:face:b00c:0:16

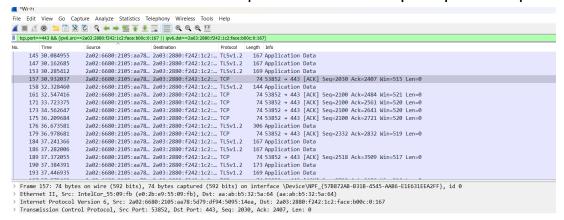


Message Sizes over Time for Group: 2a03:2880:f242:1c2:face:b00c:0:167--2a02:6680:2105:aa78:5d79:df94:5095:14e



חלק מקובץ הטקסט שמכיל את סיכום הפאקטות שנתפסו:

שמנו לב שחוזרות על עצמן פאקטות בגודל 74 בייט והן רעש מבחינתנו. לכן כרגע אנו לא מחשיבות אותן כחבילות שמעבירות הודעות ממש אלא כרעש קבוע של האפליקציה. כלומר לחיצות הידיים , ואכן רואים בהקלטת הווירשארק המתאימה(למטה) שכל הפאקטות שגודלן 74 הן מסוג ACK- התחלת קשר.



ייצאנו לקובץ CSV את הפאקטות המתאימות (ללא הרעש) בעזרת סינון נוסף בווירשארק.

ביצענו ממוצע על עמודת הגודל.

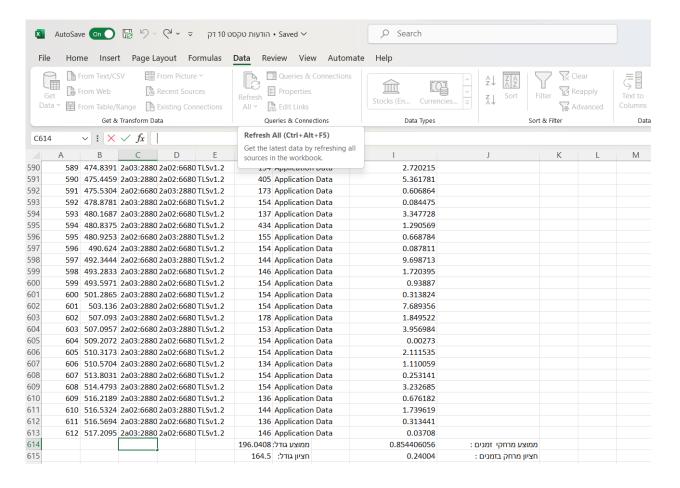
גילינו שהגודל הממוצע הוא 196 עבור כל פאקטה.

החציון עבור גודל הפאקטות הוא 164.5

ממוצע המרחק בין חבילה לחבילה הוא :0.8544

והחציון הוא :0.24

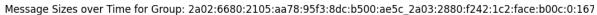
כפי שניתן לראות בצילום המסך של האקסל:

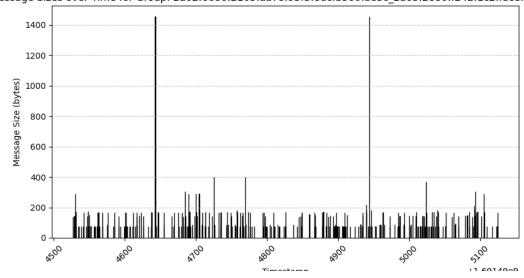


האזנה מספר 2- תמונות:

. הקלטנו 10 דקות של שליחת תמונות

: ואלו היו גרפי התוצאות

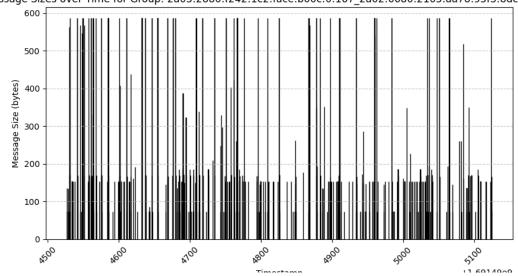




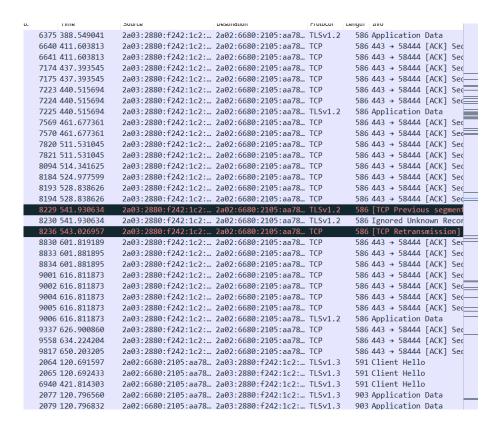
<u>הגרף העליון:</u> מייצג את מה שיצא מהמחשב שלנו אל שרתי ווצאפ. אולם אנחנו שלחנו את התמונות ממכשירים אחרים אל הקבוצה. לכן הוא לא משמעותי כרגע למעקב אחר שליחת תמונות

<u>הגרף התחתון:</u> מייצג מבחינתנו את התעבורה המרובה בתמונות שנשלחה ועברה משרתי ווצאפ אל המחשב שלנו.

Message Sizes over Time for Group: 2a03:2880:f242:1c2:face:b00c:0:167_2a02:6680:2105:aa78:95f3:8dc:b500:ae5c



בהקלטת הווירשארק שמנו לב כי הרבה מהפאקטות באורך 590 בערך הן מסוג ACK ולכן רעש מבחינתנו



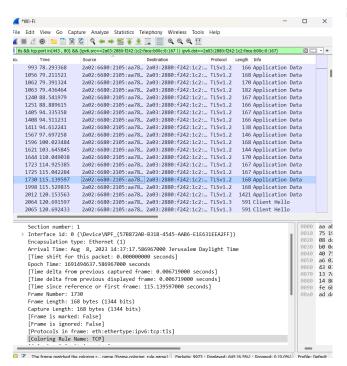
בשני הגרפים נתמקד לרגע במשותף -קבוצות של עמודות צפופות בגובה באזור 150-180 .

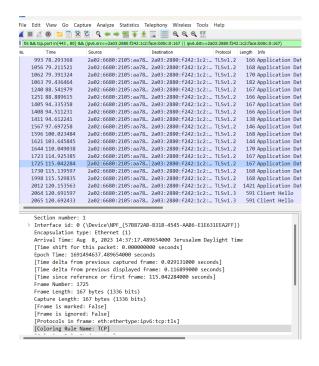
: נביט בהקלטת הווירשארק במיון לפי מקור השולח

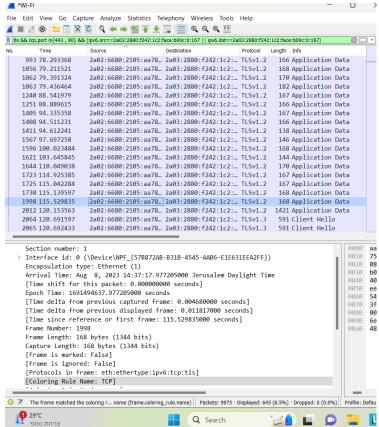
נבחין בצמדים או שלישיות של חבילות שהגיעו בזמן ממש דומה ברמת הבדל של 0.01

שניות. ההנחה שלנו היא שכל צמד כזה הוא תמונה שהמידע בה חולק כמה חבילות. למשל :

אפשר לראות בזמן 115...(הפאקטות המסומנות בכחול) :הבדלים מזעריים בזמן השליחה בשעה 14:37 ותכף נסתכל בקבוצה מה נשלח בשעה זו.

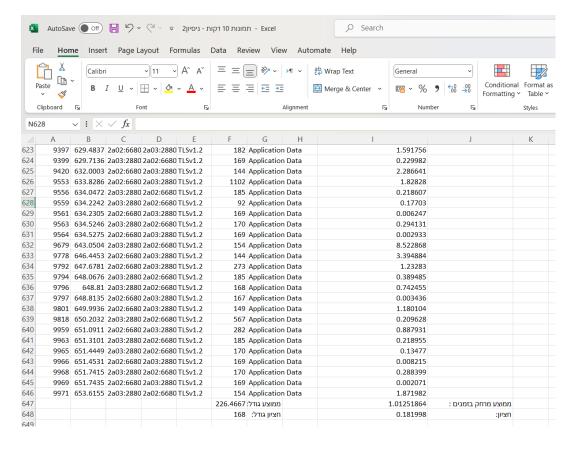






ואכן בשעה 14:37 נשלחה תמונה והיא כנראה חולקה ל3 חבילות כל אחת 167/8 בגודלה.





בקובץ האקסל שייצאנו מהוויירשארק לאחר מכן ביצענו median וגם מער average גודל הפאקטות והמרחק בזמנים בין פאקטה לפאקטה. כמו שמובא בצילום:

אכן הגודל המדיאני היה 167 . מה שתומך בהשערתנו שהחלוקה היא לפאקטות בגודל 167 בערך.

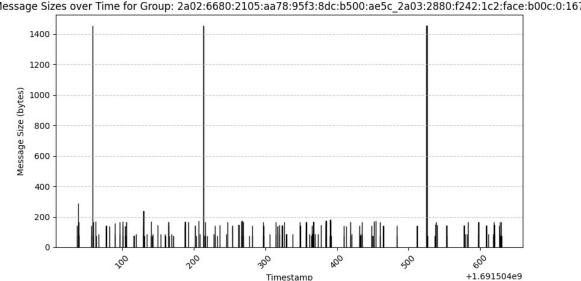
ממוצע מרחק בזמנים : 1.01251864

חציון מרחק בזמנים: 0.181998

כלומר חציון המרחק בזמנים קטן יותר מהודעות טקסט. ואילו הגודל של חבילה ממוצעת דומה.

כלומר הגדלים הם דומים אבל נשלחים בצמדים צפופים יותר, כי הקבצים גדולים יותר ומחולקים ליותר חתיכות.

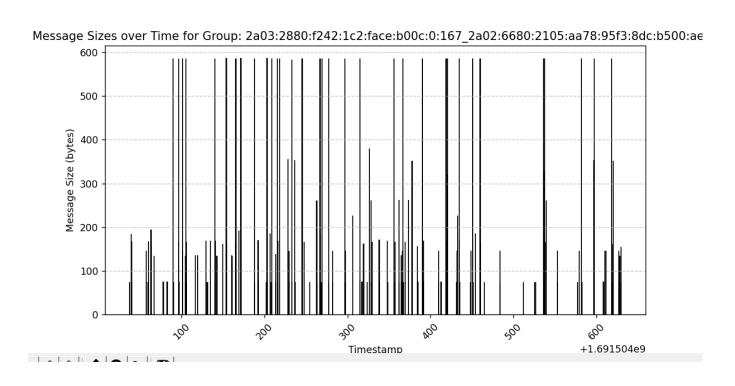
האזנה מספר 3 - סרטונים:



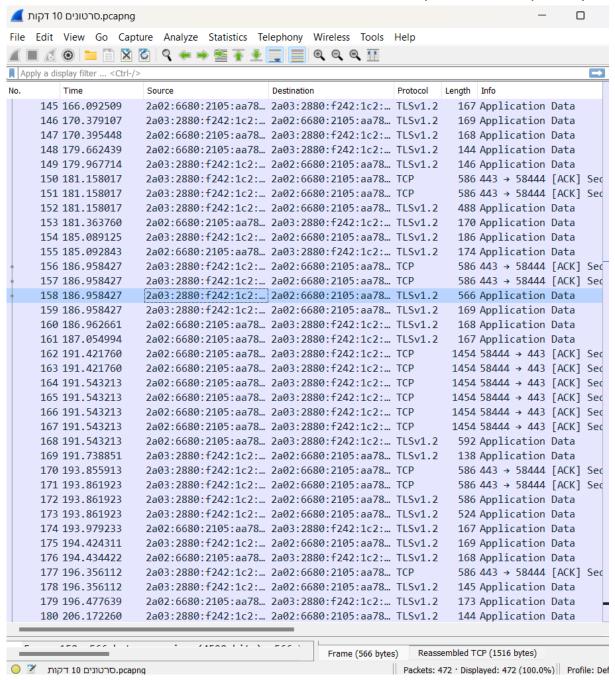
Message Sizes over Time for Group: 2a02:6680:2105:aa78:95f3:8dc:b500:ae5c_2a03:2880:f242:1c2:face:b00c:0:167

הגרף העליון: מייצג את מה שיצא מהמחשב שלנו אל שרתי ווצאפ. אולם אנחנו שלחנו את התמונות ממכשירים אחרים אל הקבוצה. לכן הוא לא משמעותי כרגע למעקב אחר שליחת תמונות

<u>הגרף התחתון:</u> מייצג מבחינתנו את התעבורה המרובה בסרטונים שנשלחה ועברה משרתי ווצאפ אל המחשב שלנו.

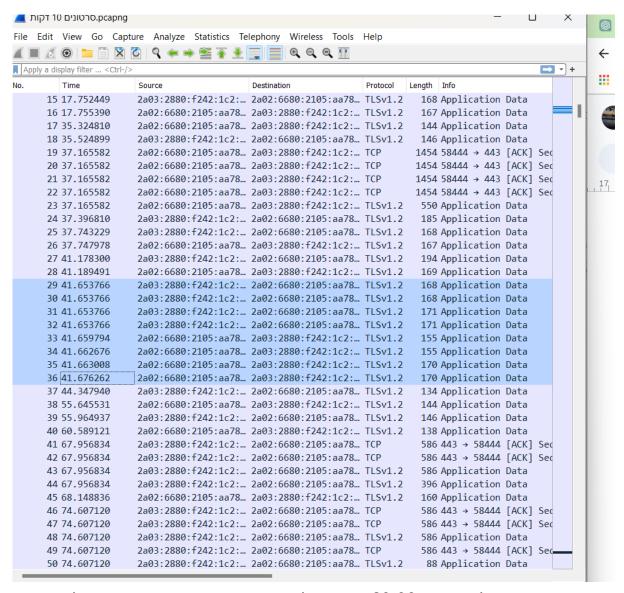


: נבחן את הקלטת הווירשארק המתאימה



ניתן לראות שהעמודות הגבוהות מציינות בכל פעם קבוצת פאקטות גדולות (הקוד מצוות פאקטות שנשלחות בהפרש זמן קטן מ0.1 שניות.) למשל פאקטות 162-167 נמצאות בעמודה אחת בגרף העליון. בנוסף ניתן להבחין שהפאקטות בסוג של רצף לאחר יצירת קשר ACK ובכל פעם הגודל של החחבילות המחולקות משתנה...כנראה לפי גודל הסרטון. למשל פאקטות 170-172 ככל הנראה מיייצגות תחילת קשר ואחריו שליחת סרטון בחתיכות בגדלים 586 +524 אחכ נראה שינוי קל בזמנים ומעבר להודעה אחרת גם פאקטות 175 ו176 נראות משוייכות להודעה אחת , ייתכן וזהו לא סרטון אלא הודעה בגודל קטן יותר שהתפצלה לגדלים 169+168

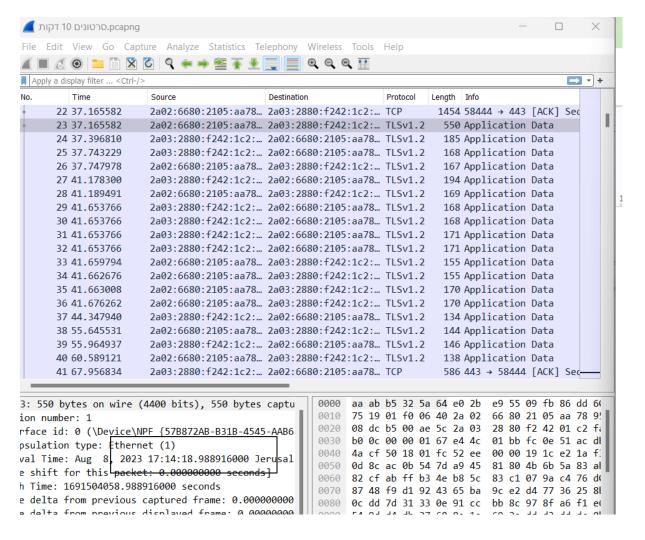
: רצף נוסף מענייין שמצאנו



הרצף המסומן בכחול פאקטות 29-36 נראות כולן בהפרשי זמן ממש קטנים כאשר לפי המידע שבתוך הפאקטות הן שייכות לזמן : :17:14:23

לפי המידע גם ראינו שהפאקטות שלפניהן ואחריהן שייכות כולם לדקה הזו , יתר על כן כל ACK בפאקטות בין ACK בפאקטה 22 ועד ה

לכן נצרף צילום מסך חדש:

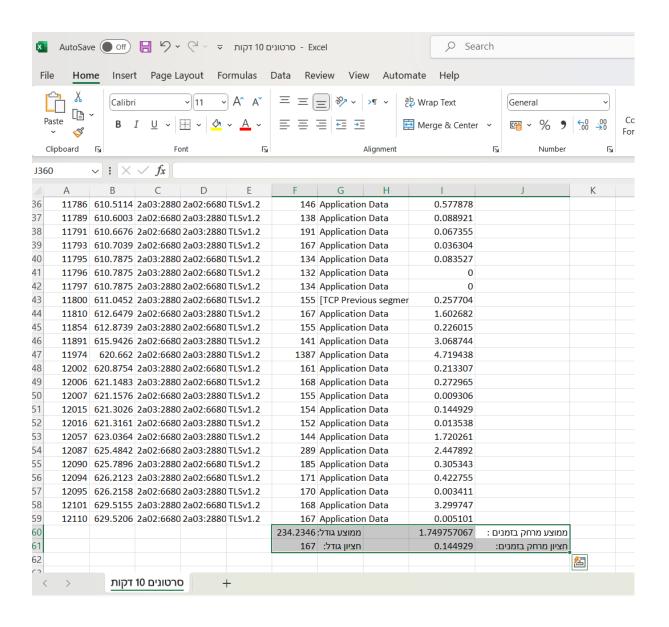


ואכן בקבוצת הווצאפ נשלחו שלושה סרטונים בדקה הזו שניים ממקור שלנו לשרת ווצאפ ואחד שהתקבל משרת ווצאפ אלינו.

היות ואנו לא מבחינות בוודאות בהבדל בין הסרטונים כי הפאקטות מעורבבות נסתפק בכך שרואים שהפאקטות מחולקות לחתיכות קטנות בגדלים 146-185 והראשונה שאחרי יצירת הקשר גדולה יותר (586)

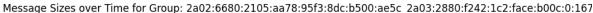
אף על פי כן גם כאן כשביצענו ממוצע וחציון על גודל הפאקטות לא קיבלנו הבדל משמעותי בין ההאזנות כלומר המממוצע היה 234 והחציון - 167

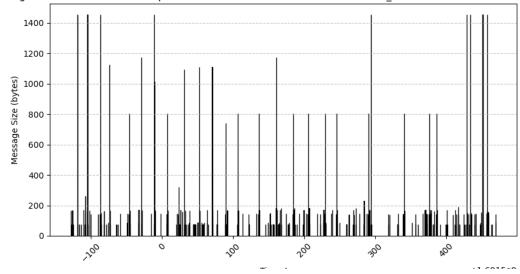
> כאן הממוצע מרחק בזמנים גדל ל 1.749 ואילו החציון מרחק בזמנים הוא 0.144 - קטן יותר מהתמונות



האזנה מספר 4 - הודעות קוליות:

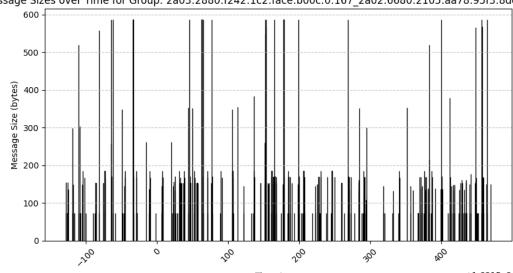
: האזנו במשך 10 דקות להקלטות והודעות קוליות ואלו הגרפים



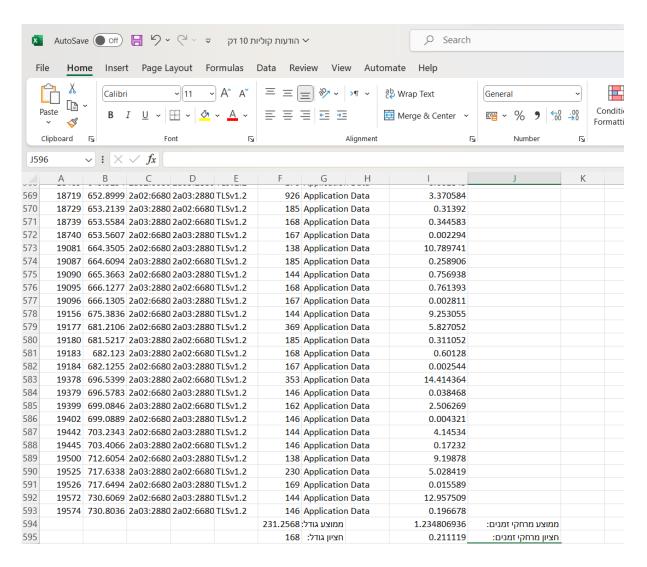


הגרף העליון: מיצג את ההקלטות שנשלחו מהמכשיר שלנו אל שרתי ווצאפ. הגרף התחתון : מייצג את התעבורה מרובת ההקלטות שנשלחה משרתי ווצאפ אלינו.

Message Sizes over Time for Group: 2a03:2880:f242:1c2:face:b00c:0:167_2a02:6680:2105:aa78:95f3:8dc:b500:ae5c



גם כאן לפי האקסל החציון הוא הודעה בגודל 168 בדומה לסרטונים ולתמונות ההשערה שלנו כרגע היא שהשרת של ווצאפ מחלק את ההודעות לחתיכות בגודל הזה כי ניתן להבחין ברצפים של הודעות בזמנים צפופים נבחין בקובץ האקסל שבו אספנו את הפאקטות :



כאן ממוצע מרחקי הזמנים בין הפאקטות היה 1.234

והחציון היה 0.2111- יותר גדול מהסרטונים והתמונות אבל יותר קטן מהודעות טקסט.

מסקנות המחקר:

האם ניתן להסיק מסקנות על הקבוצות בהן לקחנו חלק באמצעות הטכניקות המפורטות במאמר?

בהתחשב בשני מקרים: 1. התוקף נמצא רק בקבוצה אחת 2. התוקף נמצא בכמה קבוצות ווצאפ בו זמנית.

מטרתן של הטכניקות המפורטות במאמר:

- להצליח לזהות "אירועים" ע"י בחינה של צפיפיות זמן שליחת הפאקטות ויכולת לצרף מספר פאקטות לכדי מסר אחד - אירוע אחד.
 - להתחקות אחר מנהלים ומשתתפים בקבוצות, דרך מעקב והשוואת צורת התעבורה או צפיפות אירועים בין משתמש ספציפי לתעבורת הקבוצה.
 - להתחקות אחר התעבורה המרכזית שמועברת בקבוצה.

אנחנו השתדלנו לבצע גרפים שיתארו לנו את תעבורת הקבוצה כמו שמתואר במאמר. וזו הכנה טובה לשימוש בשיטותיהן. כלומר הקוד שיצרנו קולט את הפאקטות שעוברות ברשת, ומחלק לפי תקשורת בין שרתים את הודעות. הוא מייצר גרף עבור כל ערוץ תקשורת ובתוכו מסווג פאקטות "קרובות מספיק" לכדי הבנה שזוהי "הודעה" אחת.

ע"מ להגיע למסקנות על מנהלים ומשתתפים פעילים בקבוצה נצטרך השוואה בין תעבורת משתמש ספציפי בקבוצה לגרפי התעבורה שמצאנו. אם נמצא דמיון מספיק חזק לפי האלגוריתמים שמתוארים במאמר- נוכל לדעת שאותה זהות שאנו עוקבים אחריה פעילה בקבוצות.

כדי להשיג תעבורה של משתמש ספציפי נצטרך אישור להאזנה כמו שציינו בסיכום המאמר למעלה, ורק אז נוכל להאזין לתקשורת פרטית של משתמש.

ע"מ להתחקות אחר סוג התעבורה בקבוצה , נצטרך להתבונן עמוקות בהקלטות התעבורה ולנסות לפענח האם צפיפות ההודעות מראה על קבצים גדולים מחולקים או על הודעות טקסט במרווחים שונים. (מה שהקוד שלנו מבצע.)

נשים לב להבדל בין שני המקרים שהזכרנו לעיל:

- התוקף נמצא בקבוצה בודדת מה שאנחנו הדגמנו במחקר, בכל פעם יצאנו מנקודת הנחה שהתעבורה היא רק מקבוצה אחת אליה האזנו.
 אזי אפשר להיות בטוחים שכל התעבורה מהקבוצה הזו בלבד. במקרה כזה יהיה קל יותר להשוות את התעבורה של משתמש בודד אל מול התעבורה בקבוצה.
- התוקף נמצא בכמה קבוצות אזי קשה לדעת מאיזה קבוצה הגיעה כל הודעה (בגלל ההצפנה של ווצאפ) אנו רק מצליחים לראות שזה הגיע מהשרת של ווצאפ ואז קשה יותר לסווג איזה הודעה הגיעה מאיזה קבוצה. במקרה כזה ככל הנראה ההתאמה בין תעבורה של משתמש בודד לתעבורה של קבוצה ספציפית תהיה קשה הרבה יותר.

על מנת להשתמש בהאזנה לנתוני הקבוצה והצלבת נתונים לגילוי משתמש ווצאפ החשוד במעורבות בקבוצה , נעקוב אחרי התקשורת של החשוד ונרצה לבדוק - א. האם הוא משתתף בקבוצה, ב. האם הוא פעיל בקבוצה.

אם הוא נמצא בקבוצה בודדת אז הבדיקה תהיה פשוטה.

נשווה את כל תעבורת הרשת משרת הווצאפ אליו עם תעבורת הרשת המתקבלת אצלנו בזמן שאנחנו בשקט בקבוצה (או אם אנחנו לא בשקט נצטרך להוסיף את ההודעות שאנחנו שלחנו). אם ניכרת התאמה באחוז גבוה נדע שהחשוד משתתף בקבוצה.

לגבי מעורבות בקבוצה הבדיקה תהיה דומה. נשווה את ההודעות היוצאות מהחשוד אל שרת ווצאפ עם הפעילות הכללית בקבוצה כפי שאנחנו רואים מההאזנה שלנו אליה, ונראה כמה התאמה יש בזמני שליחת ההודעות ובגדלים שלהן.

במקרה בו החשוד פעיל בכמה קבוצות בו זמנית יהיו לנו הודעות לא רלוונטיות שיפריעו לנו בהתאמת הנתונים, אך עדיין אפשרי לזהות מעורבות בקבוצה.