# פרויקט גמר רשתות תקשורת

מאיה זנד- 318964699 | אופיר שטרית-324249150
קישור לפרויקט: https://github.com/ofirshitrit/Networks-Project

**חלק ראשון – מאמר:**

הרעיון המרכזי של המאמר הוא להדגים כיצד שירותי מסרים מיידיים (IM) יכולים להיות פגיעים להתקפות ניתוח תעבורה, מה שיכול לאפשר ליריבים להשיג מידע רגיש על משתמשים. המאמר מציע התקפת ניתוח תעבורה מעשית שניתן להשתמש בה כדי לזהות את חברי ערוץ היעד בשירותי מסרים מיידיים פופולריים כמו Telegram, Signal, WhatsApp, Viber.

כדי לבצע את ההתקפה, התוקף צריך קודם כל להשיג מידע אמין לגבי התנועה של ערוץ היעד. ניתן לעשות זאת בשלוש דרכים:

(1) על ידי הצטרפות לערוץ היעד כחבר והקלטת ההודעות שנשלחו בערוץ יחד עם המטא דאטה שלהן.

(2) על ידי פרסום הודעות לערוץ היעד אם התוקף הצטרף לערוץ והוא מסוגל לפרסם הודעות.

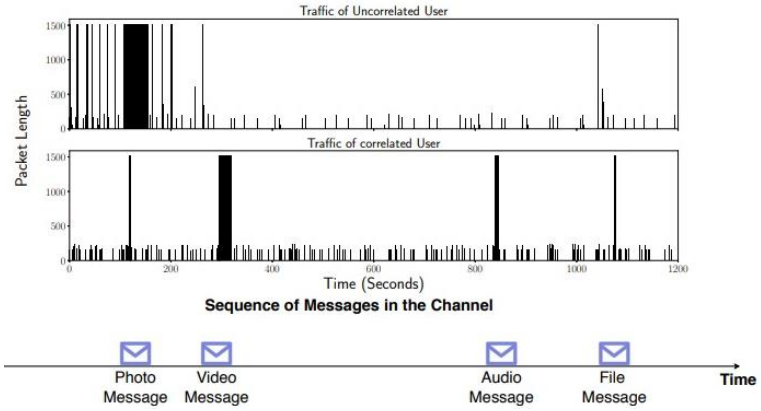
(3) על ידי השגת תפקיד מנהל לערוץ היעד.

לאחר שהתוקף קיבל מידע אמין אודות התנועה של ערוץ היעד, הוא יכול להשתמש באלגוריתם זיהוי כדי להתאים את דפוסי התנועה של המשתמשים לדפוסי התנועה אודות ערוץ היעד. המאמר מציג תוצאות ניסוי המדגימות את יעילות המתקפה המוצעת.

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 |

טבלה 2 מציגה עבור כל סוג הודעה (טקסט/תמונה/וידיאו וכו') את כמות ההודעות, גודלה הממוצע של ההודעה לפי סוג, טווח גדלי ההודעה וכו'. מטבלה זו ניתן להסיק על סוגי ההודעות הנפוצות ביותר ואלו שנמצאות בשימוש תדיר ביותר אצל משתמשים.




איור 8 מציג את ביצועי המתקפה המוצעת על שירותי מסרים מיידיים שונים, ומדגים כי המתקפה יעילה במגוון פלטפורמות. התוצאות מראות שניתן להשתמש במתקפה המוצעת כדי לזהות את חברי ערוץ היעד ב-Signal, WhatsApp ו-Viber ברמת דיוק גבוהה, גם כשהערוץ מוצפן.

המאמר מגיע למסקנה כי המתקפה המוצעת היא דרך מעשית ויעילה לבצע התקפות ניתוח תעבורה על שירותי מסרים מיידיים וכי המשתמשים צריכים להיות מודעים לסיכונים הפוטנציאליים הכרוכים בשימוש בשירותים אלו.

## חלק שני – מימושים ותובנות:

הסבר על הפרויקט: בפרויקט זה התבקשנו להקליט תקשורת בקבוצות IM (Instant messeging). אנו בחרנו לקיים תקשורת באפליקציית וואצאפ-ווב והקלטנו אותה בווירשארק. הודעות אלו סווגו לסוגים שונים: הודעות, תמונות, סרטונים, הקלטות וכדומה, כאשר ביצענו הפרדה בין סוגי ההודעות השונים לפי קבוצות.
נרצה לבדוק האם ניתן לסווג כל אחת מן הקבוצות לפי זמן וגודל הפאקטות שהקלטנו בווירשארק.

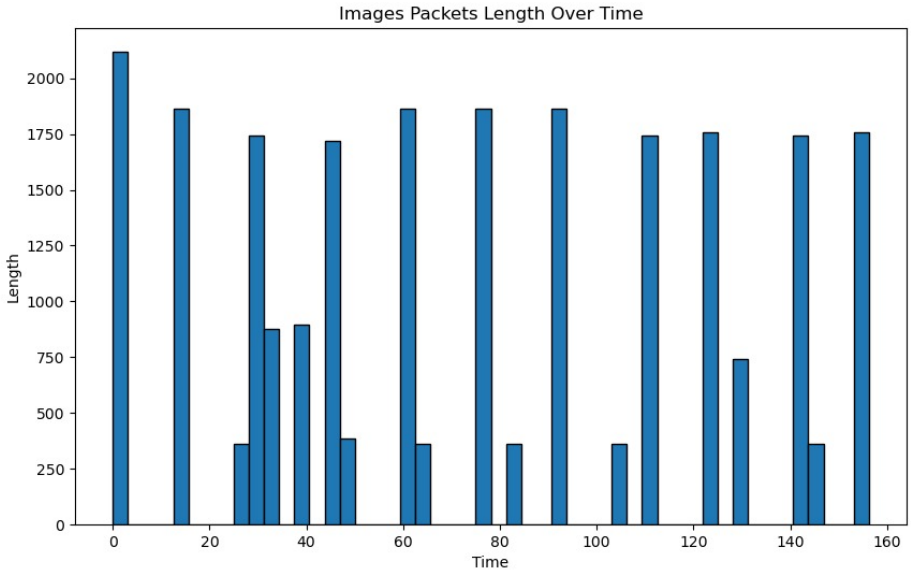
וואצאפ ווירשארק: מאחר וההודעות באפליקציית וואצאפ הן מוצפנות, לא ניתן לדעת מה תוכן ההודעות. מטרתנו הייתה להבין איזה מידע בווירשארק משויך להודעות ששלחנו בקבוצות הוואצאפ השונות, על אף ההצפנה. נפרט כעת את דרך הפעולה:
בווירשארק קיימות הרבה תנועות במקביל ולכן רצינו לבצע סינון של המידע. כאשר שלחנו הודעות וואצאפ הבחנו בכך שחלק מהמידע שמוצג בווירשארק מסווג ע"י פרוטוקול TLSv1.2 שבעמודת המידע שלו מצוין שזה מידע מאפליקציה. בנוסף, הבנו שנכון לסנן גם לפי פרוטוקול TCP ב-PORT 443, ולפי source ו-destination שמותאמים למחשב שלנו ולאפליקציית וואצאפ ווב (קיבלנו את ה-IP של וואצאפ ווב באמצעות הפקודה: nslookup). ביצענו סינונים של המידע לפי שני הפרוטוקולים הללו.

פרוטוקול TLSv1.2: הוא פרוטוקול קריפטוגרפי המספק תקשורת מאובטחת על גבי רשת מחשבים. הוא משמש ליצירת חיבור מוצפן ומאומת בין שני צדדים, בדרך כלל לקוח (כמו דפדפן אינטרנט) ושרת (כמו אתר אינטרנט).

TCP PORT 443: סיננו לפי פרוטוקול זה מכיוון שהוא הפרוטוקול הסטנדרטי לתקשורת מאובטחת על גבי אתרי אינטרנט מאובטחים (HTTPS).

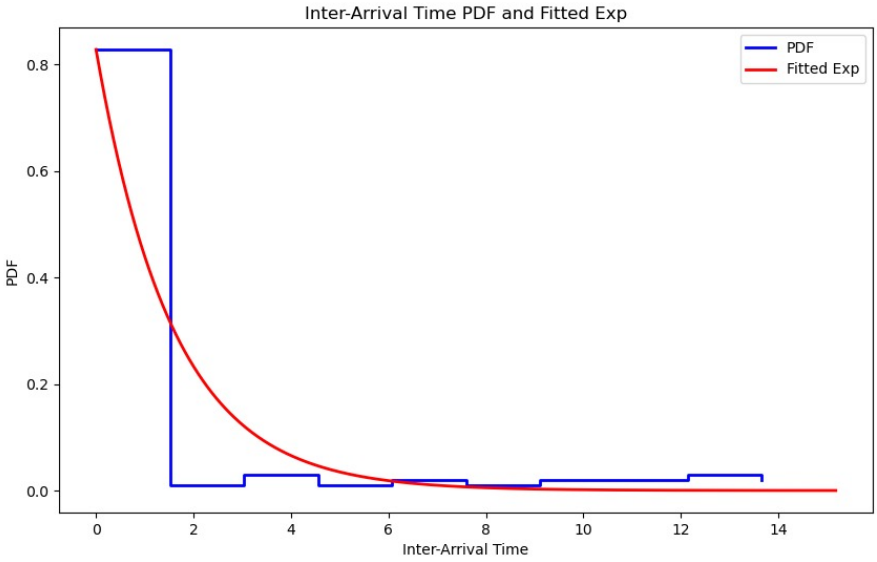
כחלק מעבודה על הפרויקט נעזרנו במאמר שעוסק בנושא של התקפות על אפליקציות של הודעות מאובטחות. במאמר זה הוצגו תוצרים שונים עבור סוגים שונים של הודעות, ואנו רצינו לבחון האם התקשורת על גבי הוואצאפ שלנו מביאה תוצרים דומים לאלו שהוצגו במאמר. נציג כעת את התוצרים שלנו עבור מידע מסוג תמונות. עבור הסוגים האחרים (וידיאו, הקלטות וכו') התוצר נראה דומה:

**גרף מספר 1:** מציג את גודל הפאקטות שנשלחו לאורך זמן:




בעת שליחת ההודעות שמרנו על קצב קבוע של כ-15 שניות בין הודעה להודעה, דבר שאיפשר לנו להבדיל בין ההודעות שאנו שלחנו לבין הודעות אחרות שנשלחו ברקע. קל לראות בגרף זה את ההפרדה הנ"ל.

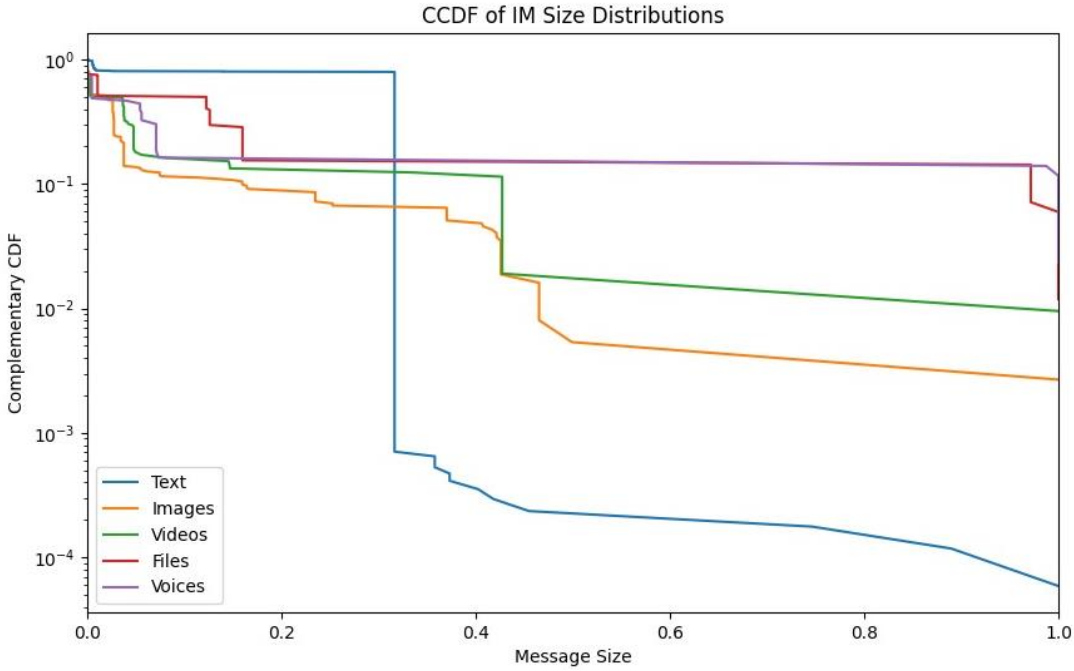
**<u>גרף מספר 2:</u>** גרף PDF הוא ייצוג גרפי של התפלגות הנתונים הרציפים לאורך טווח ספציפי של ערכים. ציר ה-X מייצג את זמני ההגעה של המידע (בשניות) וציר ה-Y מייצג את צפיפות ההסתברות של זמני ההגעה:



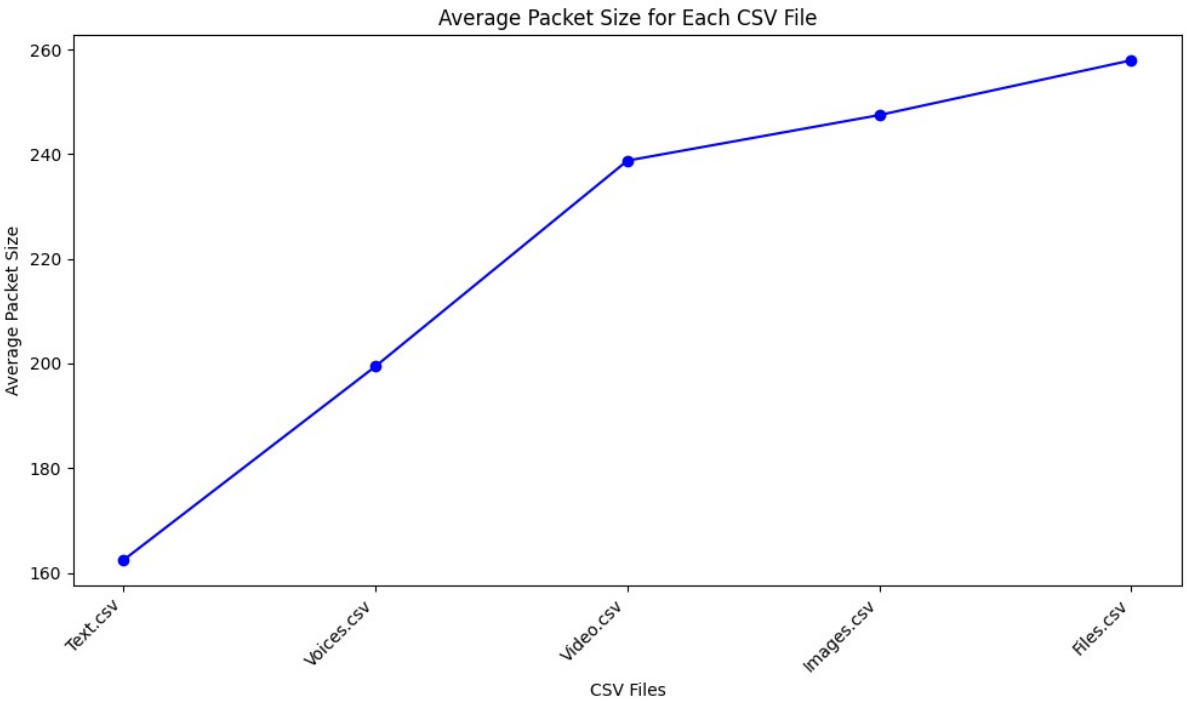

ניתן לראות בגרף שפאקטות שנשלחות במרווחי זמן נמוכים הן בעלות סבירות גבוהה להיות קשורות זו לזו.

**<u>הגרפים הבאים מתייחסים לקבוצת וואצאפ בה נשלחו כל סוגי ההודעות במקביל:</u>**

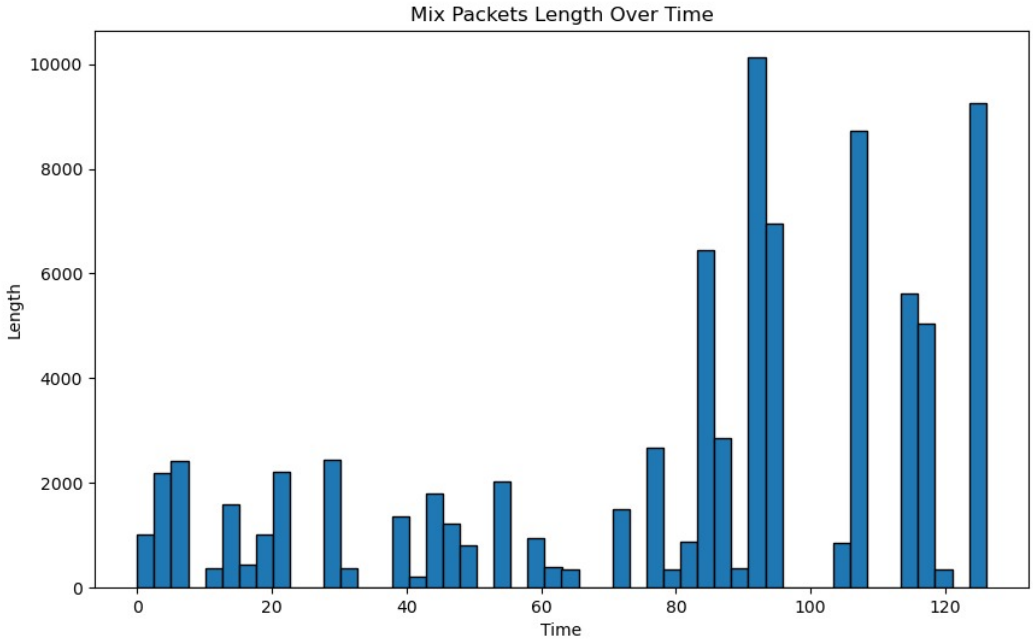
**<u>גרף 3:</u>** גרף CCDF מייצג "פונקציית הפצה מצטברת משלימה". זהו מושג סטטיסטי המשמש לניתוח והצגה של התפלגות הנתונים בדאטה. בגרף זה ציר ה-x מייצג את ערכי הנתונים (גדלי הודעות), וציר ה-y מייצג את ההסתברות שערך בדאטה גדול או שווה לציר ה-x המתאים.



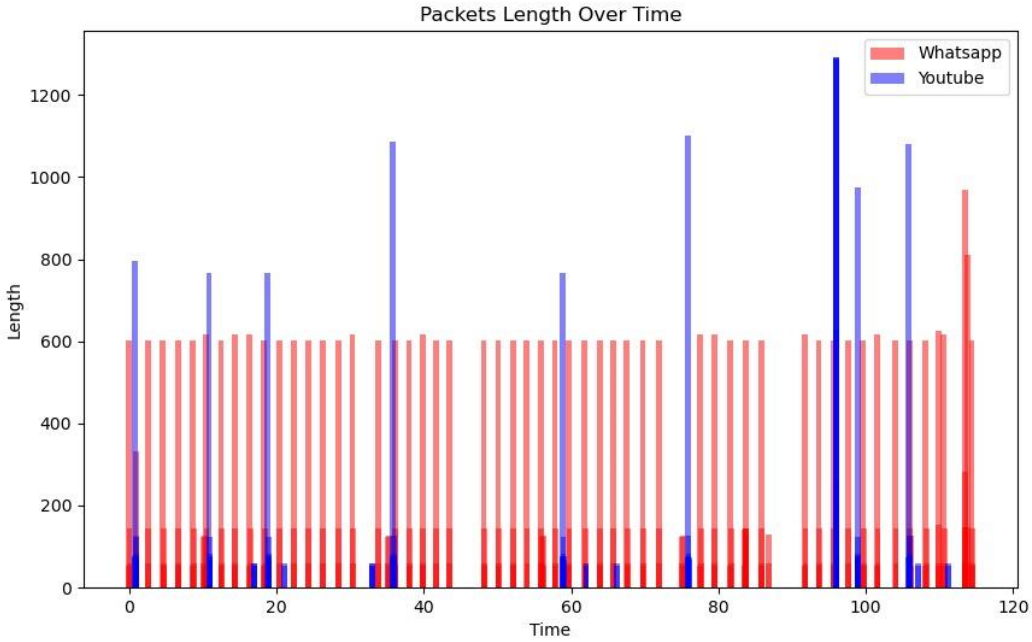
**גרף 4:** מציג את הגודל הממוצע של פאקטות מכל הסוגים של ההודעות ששלחנו:




**גרף 5:** מציג את גודל הפאקטות לאורך זמן, עבור הודעות מכל הסוגים יחד:



**גרף 6:** גרף המייצג את הפאקטות של הודעות וואצאפ שנשלחו, יחד עם פאקטות של "רעשי רקע" (במקרה שלנו- אתר יוטיוב שפועל ברקע):




נציין שבעת סיווג הפאקטות בגרף זה נתקלנו בקושי להבחין בין פאקטות הקשורות להודעות הוואצאפ שנשלחו, לבין פאקטות הקשורות לאפליקציית היוטיוב שהתנגנה ברקע. הקושי נבע מכך ששני האתרים הללו פועלים על פרוטוקולים דומים. להערכתנו הצלחנו לסווג נכון ולהבדיל את שני סוגי המידע האלו משאר המידע שראינו בתעבורה.

**סיכום:** בהתאם למחקר שהוצג במאמר ובהתאם לממצאים שהצגנו עבור התכתבויות וואצאפ שונות, ניתן לראות שעל אף הצפנת ההודעות בוואצאפ עדיין ניתן לקבל מידע שימושי ולהסיק מסקנות בנוגע להתכתבויות השונות. באמצעות סטטיסטיקות וניתוחים ניתן לזהות גדלים של פאקטות, סוגי דאטה שמועברים בהתכתבויות וכדומה. עם זאת, במאמר מצוין פתרון שיכול לשמש כאמצעי נגד התופעה, כגון הוספת תעבורת כיסוי שיכולה לפגוע ביעילות ההתקפות שמוצגות במאמר. אנו מקוות שמחקר זה יעודד את חברות ה- IM לשלב בתוכנות שלהם אמצעי נגד יעילים לערפול התנועה כדי לשמור על הצפנה מקסימלית (ולא רק של התוכן).