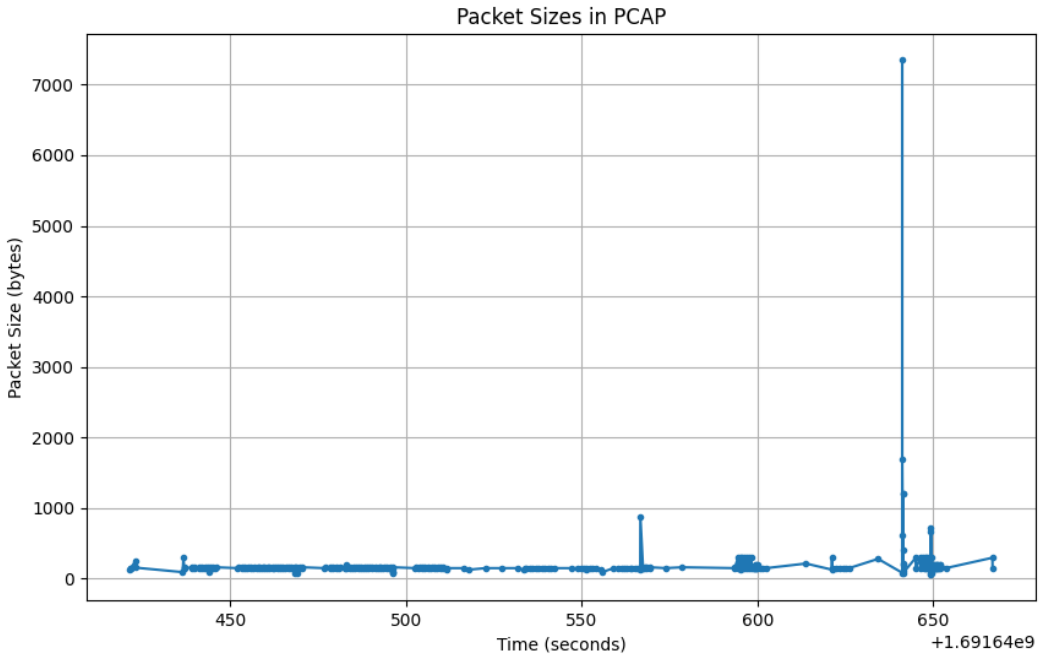
Wet Part Answers

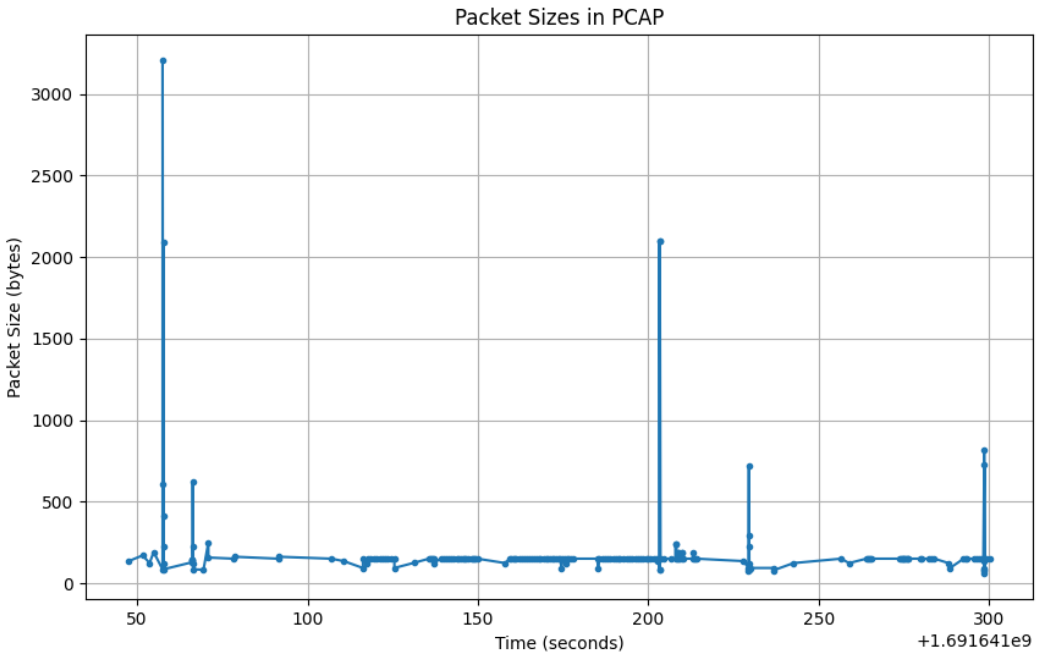
פתחנו 6 קבוצות וואטסאפ שנקראות: סרטונים, הקלטות, קבצים, תמונות, טקסט והכל מהכל. בכל קבוצה שלחנו את התיאור שלה עם קצת טקסט ביניהם.
אופן העבודה: אחד שלח את ההודעות לכל קבוצה כאשר בכל קבוצה שלחנו סוג אחר של הודעות בעוד השני הקליט את התעבורה וסינן אותה לפי:
tcp.port==433 and tls שזה אומר שאנחנו מסננים לפי פורט 433 המשמש לתעבורה מוצפנת כמו כן גם פרוטוקול tls משמש להעברת מידע מוצפן בדיוק כך מועבר המידע בווטסאפ.
כל הקלטה התבצעה בין 4-6 דקות ולאחר מכן הפסקנו את ההקלטה ושמרנו אותה.
לאחר מכן כתבנו קוד ב python שמקבל את המיקום של ההקלטה והופך אותו לגרף של גודל החבילה בבתים כתלות בזמן בשניות.
חלק זה בוצע פעמיים, פעם אחת עם הקלטות נקיות ולאחר מכן הקלטות עם רעשי רקע.

הקלטות נקיות fig 2:

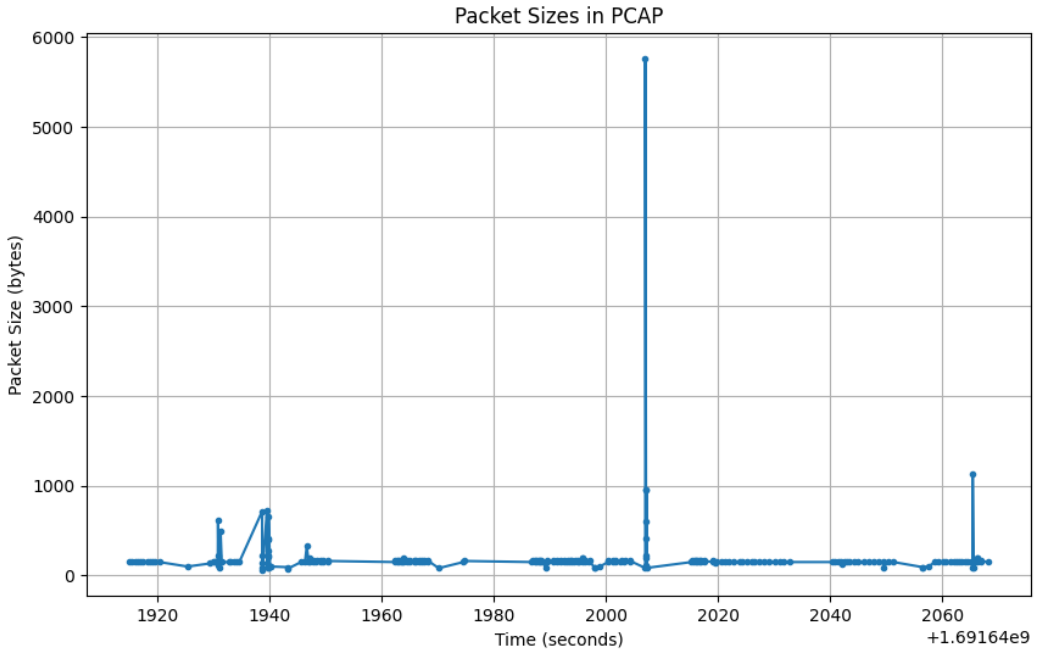
תמונות:



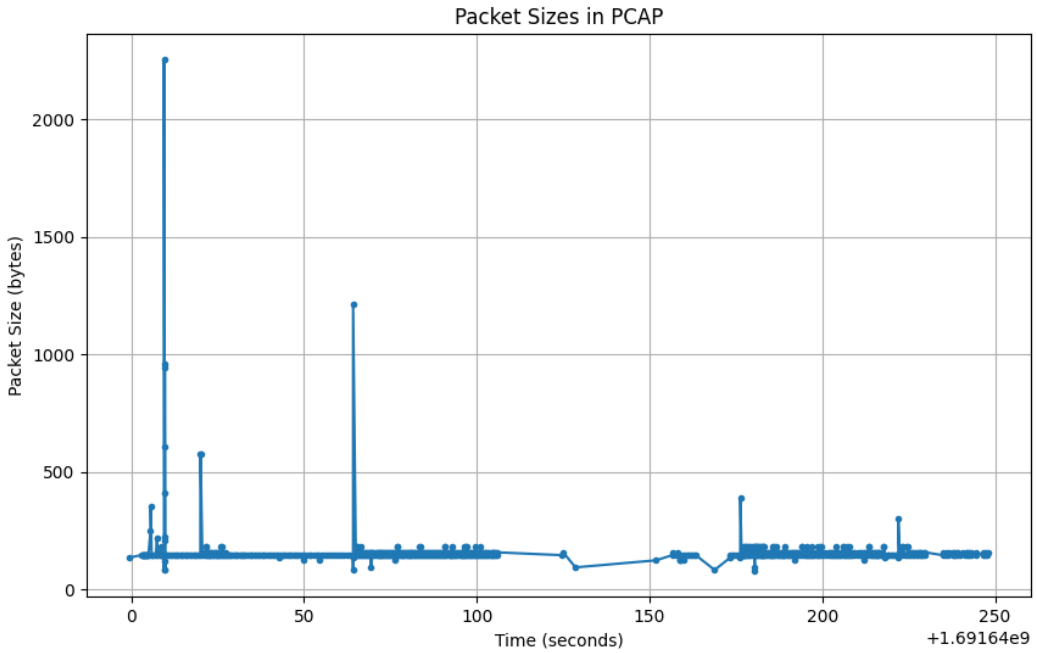
הקלטות:



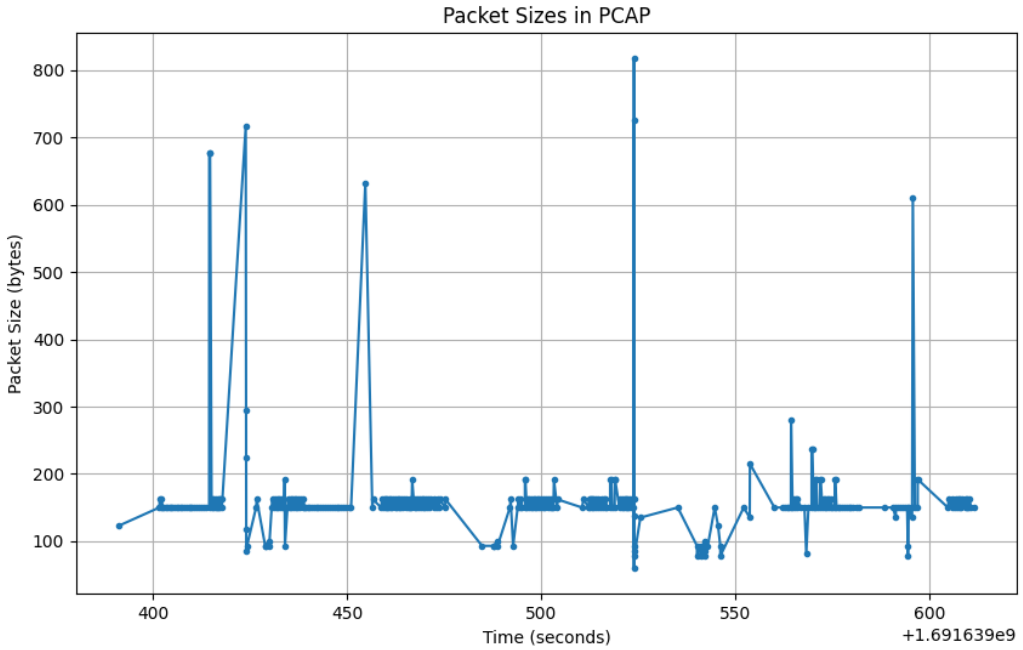
קבצים:


Packet Sizes in PCAP
Packet Size (bytes)
6000
5000
4000
3000
2000
1000
0
1920
1940
1960
1980
2000
2020
2040
2060
Time (seconds)
+1.69164e9

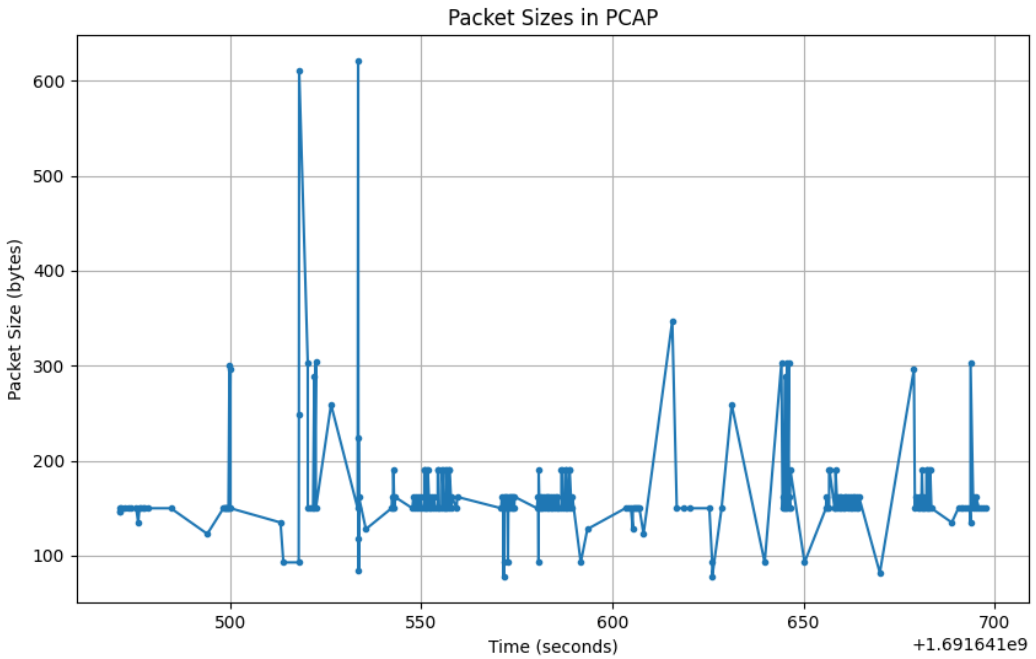
טקסט:


Packet Sizes in PCAP
Packet Size (bytes)
2000
1500
1000
500
0
0
50
100
150
200
250
Time (seconds)
+1.69164e9

וידיאו:



הכל מהכל:



הסבר על הגרפים:

all2.png (כל ההודעות ביחד)

ישנם טווחים רחבים של השהיות בין ההודעות.
גודל ההודעות משתנה באופן משמעותי, עם מספר שיאים שמצביעים על סוגי הודעות מסוימים (למשל, תמונות, סרטונים).

aud2.png (הודעות אודיו)

הודעות האודיו מראות גודל יחסית עקבי, עם שינויים קטנים.
ההשהיות בין ההודעות להודעות אודיו מתפשרות באזורים מסוימים.

file2.png (הודעות קבצים)

הודעות הקבצים מראות טווח רחב של גדלים, דבר שמצופה הואיל והקבצים יכולים להשתנות מאוד בגודלם.
ישנם טווחי השהיות שונים, המצביעים על שיתוף קבצים לא עקבי.

img2.png (הודעות תמונה)

הודעות התמונה מראות מספר אשכולות גודל מובחנים, המצביעים על כך שישנם מספר רזולוציות תמונה או דרגות דחיסה מסוימות משותפות.
ההשהיות בין התמונות הן גם הן פזורות, מה שמצביע על פעילות משתנה.

txt2.png (הודעות טקסט)

כפי שניתן היה לצפות, ההודעות הטקסט הן בגודל קטן.
ההשהיות בין ההודעות מאוד צפופות באזורים מסוימים, מה שמצביע על החלפות טקסט בתדירות גבוהה.

vid2.png (הודעות וידאו)

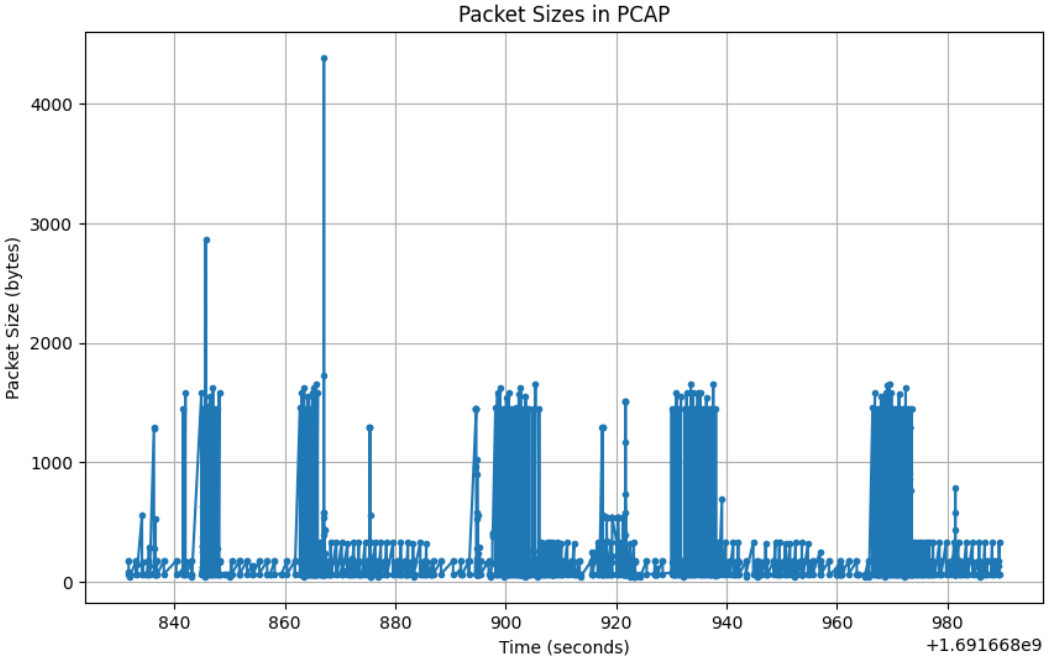
הודעות הוידאו מראות טווח רחב של גדלים, דבר שמצפה כשכוללים את ההבדלים באורך ואיכות הסרטונים.
ההשהיות בין ההודעות הן פוזרות, מה שמצביע על שיתוף סרטונים לא תדיר.

ניתוח:

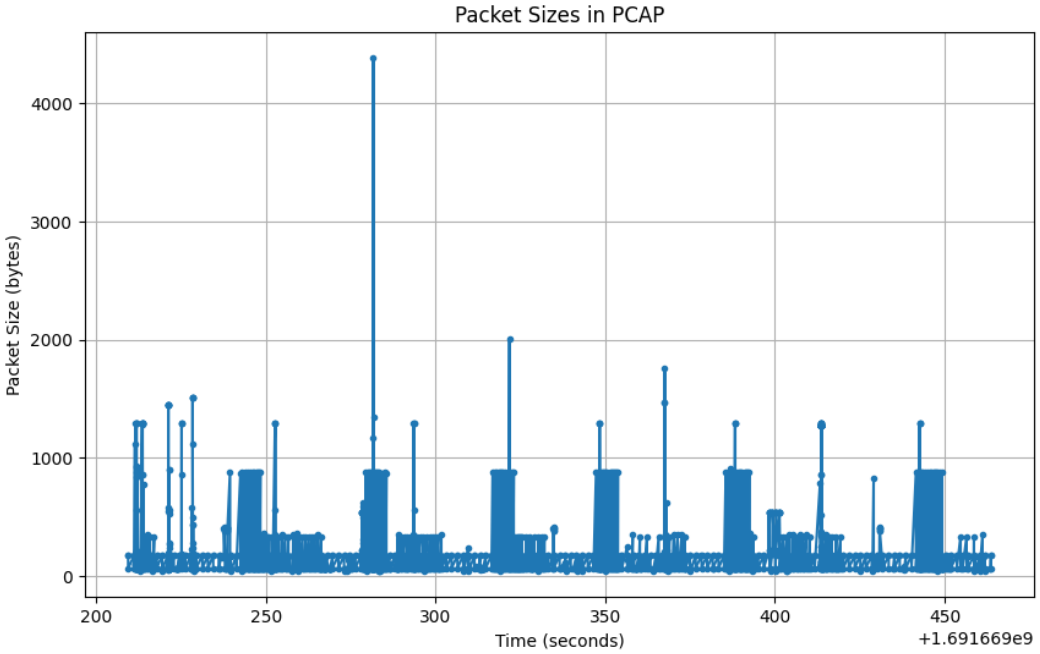
כל סוג של הודעה (אודיו, קובץ, תמונה, טקסט, וידאו) יש לו מאפיינים מובחנים בהן בגודל ההודעה ובהשהיות בין ההודעות.
הודעות טקסט, לדוגמה, מגלות החלפות תדירות (השהיות בין הודעות קצרות) וגדלים קטנים.
הודעות וידאו וקובץ, מצד שני, נוטות להיות גדולות בגודלן ומשותפות בצורה פחות תדירה.

הקלטות עם רעשי רקע fig 2:

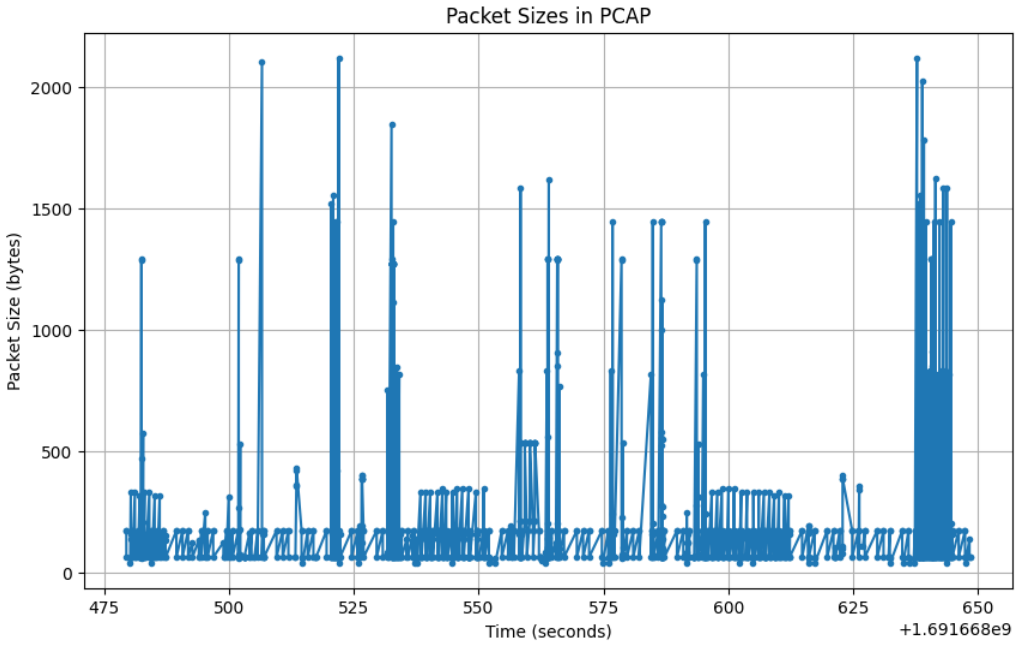
תמונות:



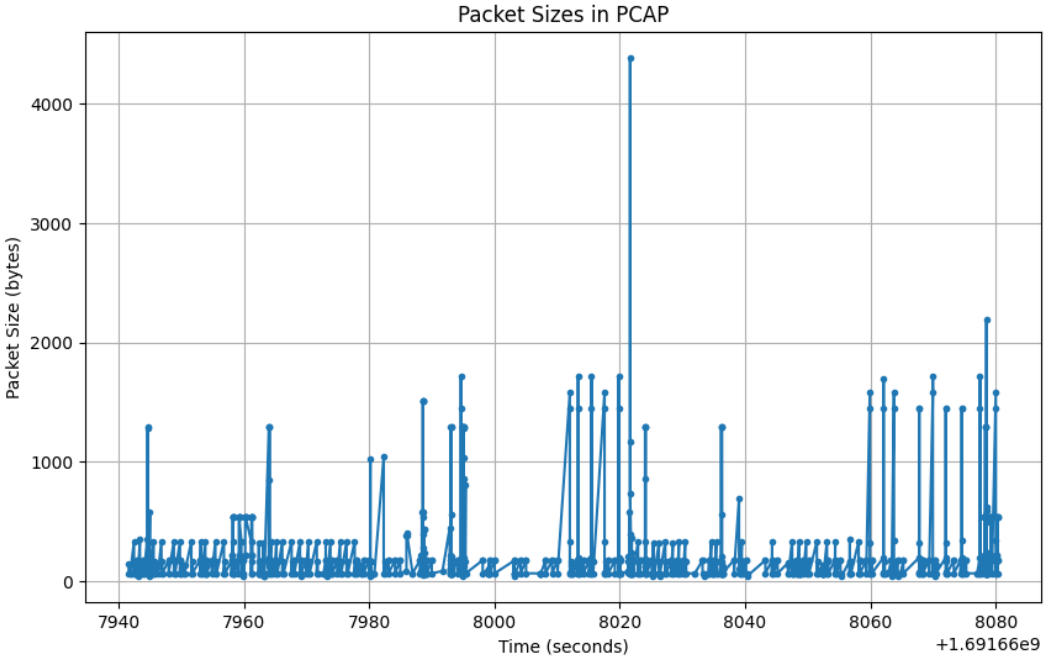
הקלטות:



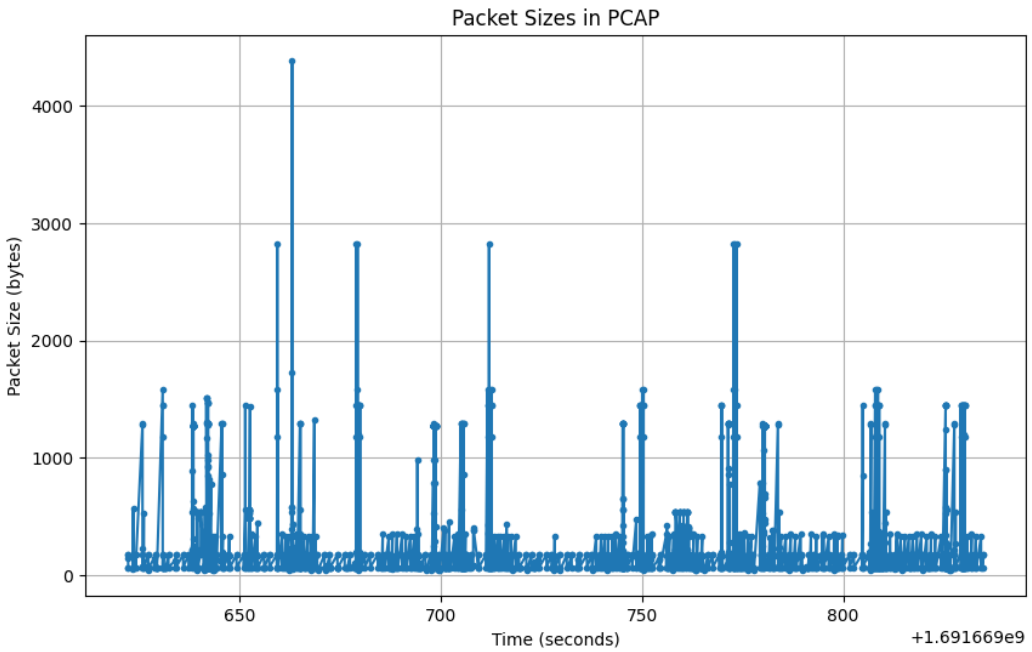
קבצים:


Packet Sizes in PCAP
Packet Size (bytes)
0
500
1000
1500
2000
475
500
525
550
575
600
625
650
Time (seconds)
+1.691668e9

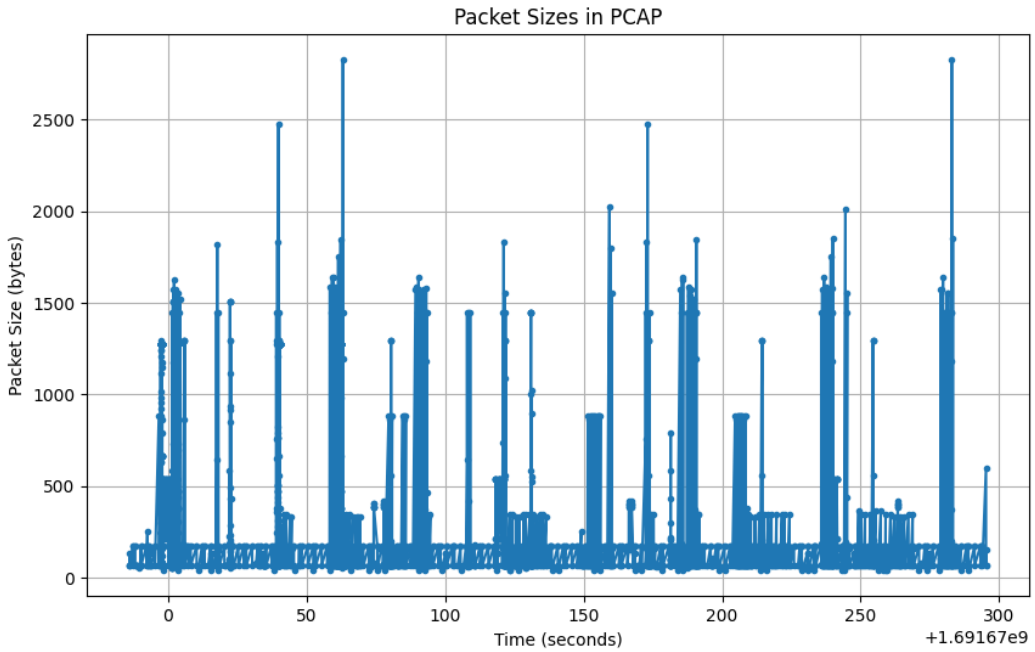
טקסט:


Packet Sizes in PCAP
Packet Size (bytes)
0
1000
2000
3000
4000
7940
7960
7980
8000
8020
8040
8060
8080
Time (seconds)
+1.69166e9

וידיאו:



הכל מהכל:



הסבר על הגרפים:

all_noise2.png (כל ההודעות ביחד עם רעשי רקע)
התרשים מצביע על התפלגות משתנה של ההשהיות בין ההודעות בהשוואה לתרשימים הקודמים, המשפיעה על הצורה בה מתקבל המידע.
ניכר שהרעש משפיע על ההתפלגות ומקשה על הזיהוי הברור של דפוסי ההודעות.

aud_noise2.png (הודעות אודיו עם רעשי רקע)
ההשהיות בין ההודעות בתרשים זה מצביעות על דפוס שונה מאוד מההודעות הקודמות.
הרעש מעוות את התפלגות ההודעות ויכול להקשות על הזיהוי המדויק של דפוס השיחה.

file_noise2.png (הודעות קבצים עם רעשי רקע)
ההשהיות בין ההודעות מצביעות על דפוס שיחה יחסית אחיד, אך עם רמת רעש גבוהה.
הרעש יכול להסתיר מסרים או להוסיף הפרעות בהזהרת המידע.

img_noise2.png (הודעות תמונה עם רעשי רקע)
בתרשים זה ניכר שההשהיות בין ההודעות ארוכות יותר מאשר בתרשימים הקודמים.
הרעש מגביר את הקושי לזהות את דפוס השיחה האמיתי.

txt_noise2.png (הודעות טקסט עם רעשי רקע)
ההשהיות בתרשים זה מצביעות על דפוס שיחה מאוד יציב.
הרעש מקשה על הזיהוי המדויק של דפוס השיחה ויכול להטעות את הזוהה.

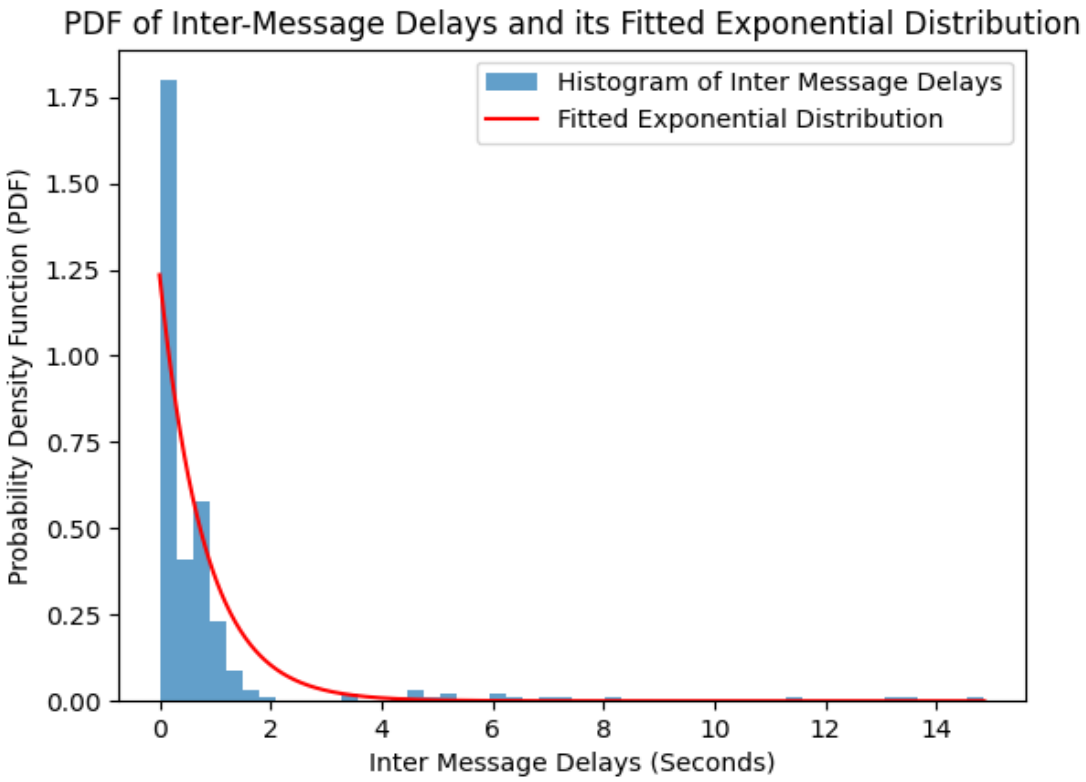
vid_noise2.png (הודעות וידאו עם רעשי רקע)
ההשהיות בין ההודעות גבוהות יותר מאשר בתרשימים הקודמים, מה שמצביע על דפוס שיחה שונה.
הרעש מעוות את ההתפלגות ויכול להקשות על הבנת הדפוס האמיתי של השיחה.

ניתוח:

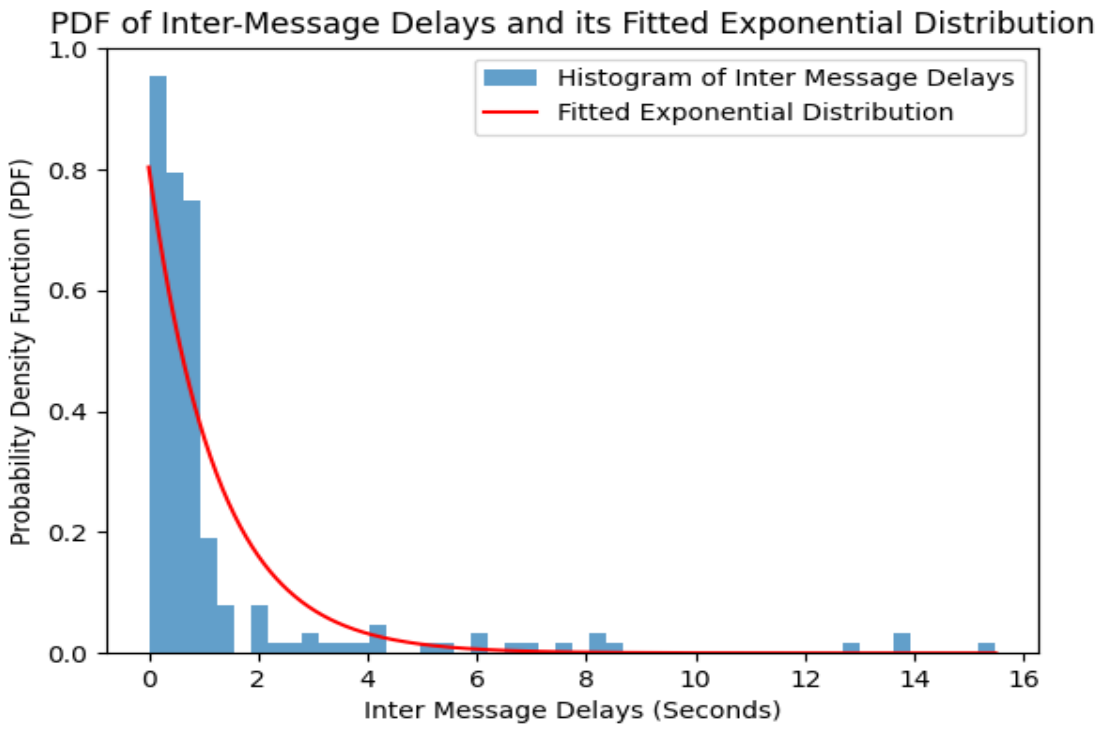
התרשימים עם רעשי הרקע מראים שינוי בהתפלגות ההודעות ובגודלם. ההשהיות בין ההודעות הן גבוהות יותר ברוב התרשימים, וכן ישנם יותר שיאים בגודל ההודעות
השינויים הללו יכולים להקשות על מתקיף אם הוא מנסה להסיק מידע מהתרשימים.
הרעש הופך את המידע לפחות צפוי ומבלבל.

הקלטות נקיות fig 3:

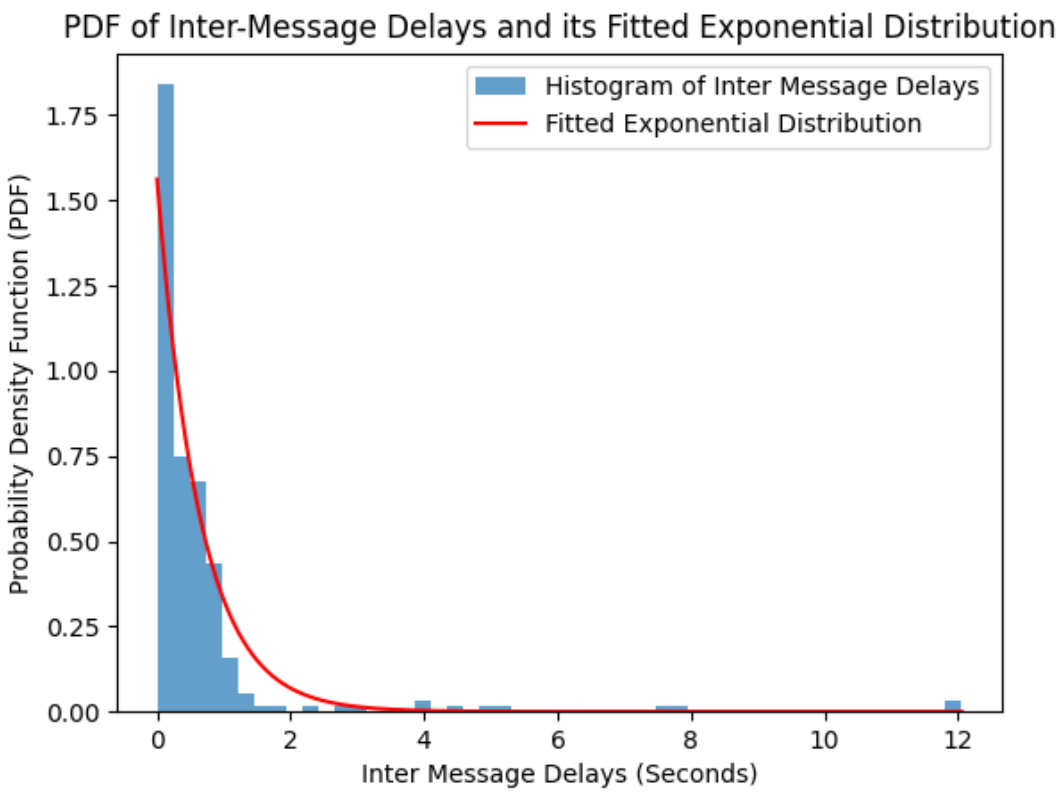
תמונות:



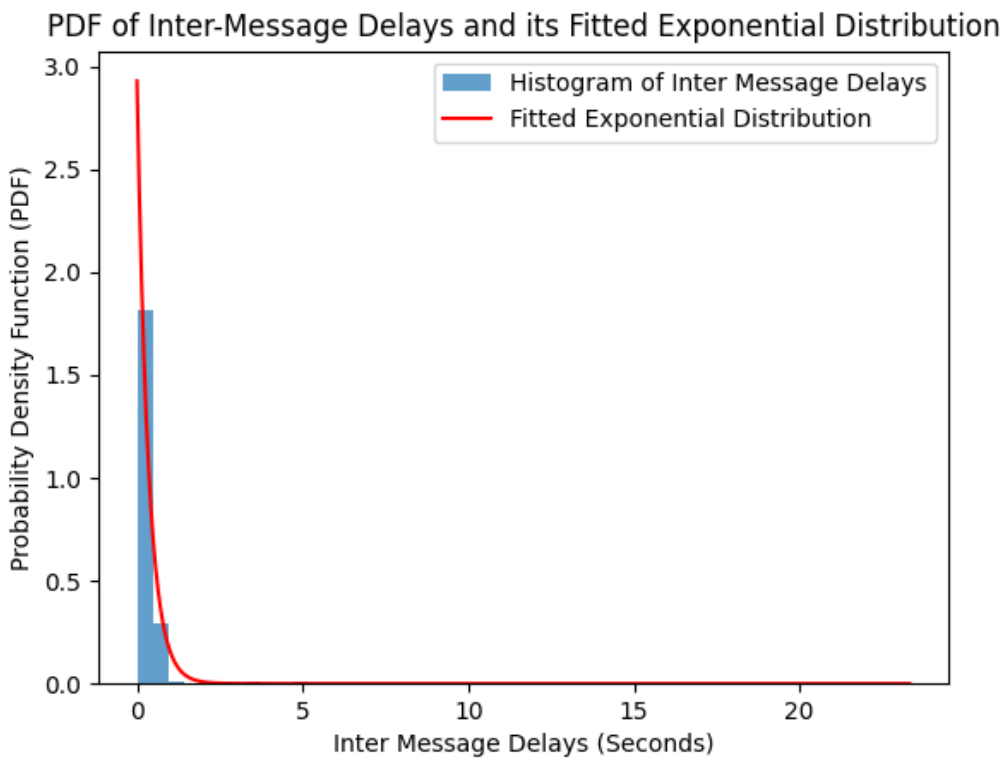
הקלטות:



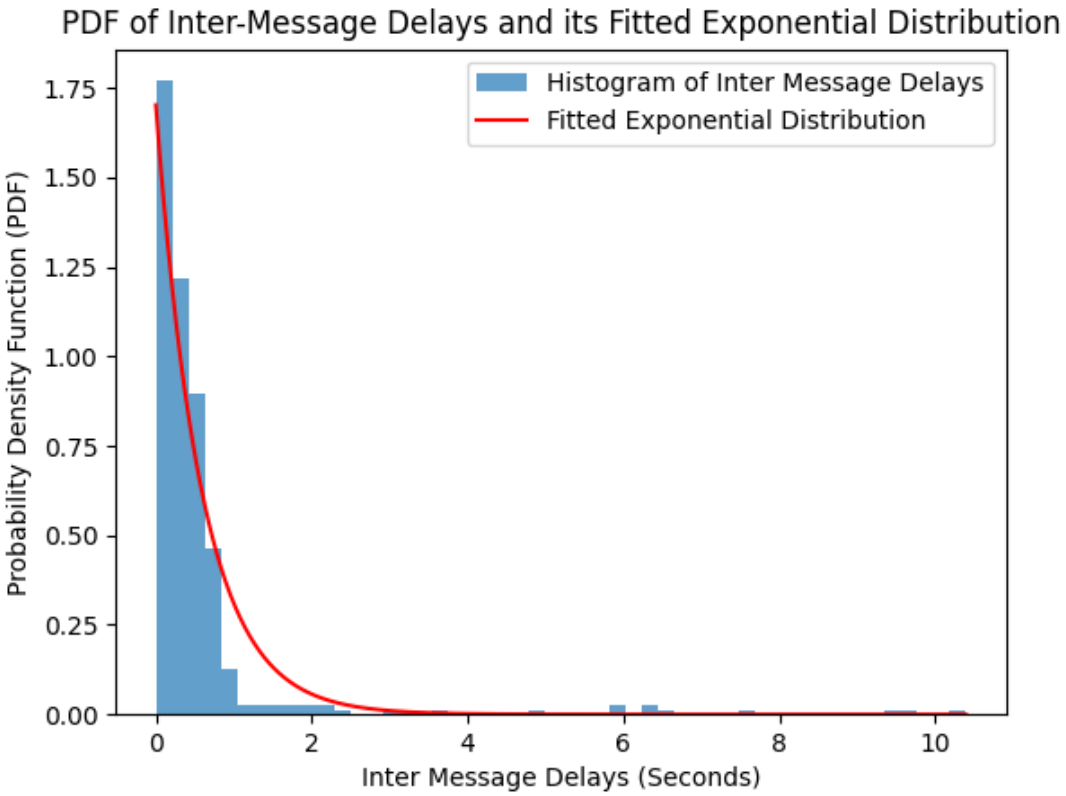
קבצים:


PDF of Inter-Message Delays and its Fitted Exponential Distribution
Histogram of Inter Message Delays
Fitted Exponential Distribution
Probability Density Function (PDF)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0
2
4
6
8
10
12
Inter Message Delays (Seconds)

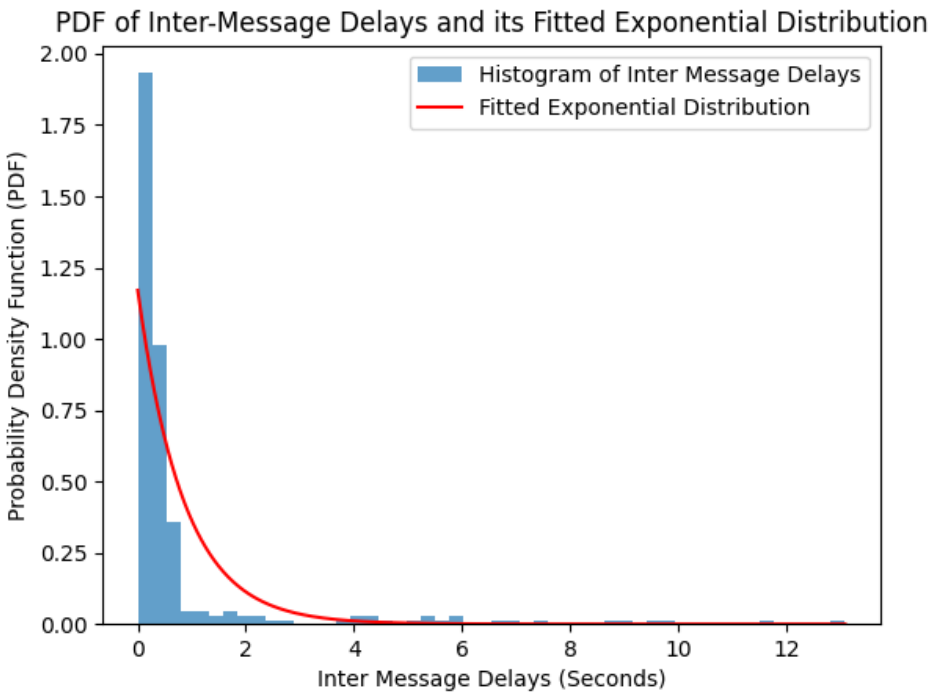
טקסט:


PDF of Inter-Message Delays and its Fitted Exponential Distribution
Histogram of Inter Message Delays
Fitted Exponential Distribution
Probability Density Function (PDF)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
5
10
15
20
Inter Message Delays (Seconds)

וידיאו:



הכל מהכל:



הסבר על הגרפים:

aud3.png, file3.png, img3.png, txt3.png, vid3.png (הודעות אודיו, קבצים, תמונה, טקסט, וידאו)

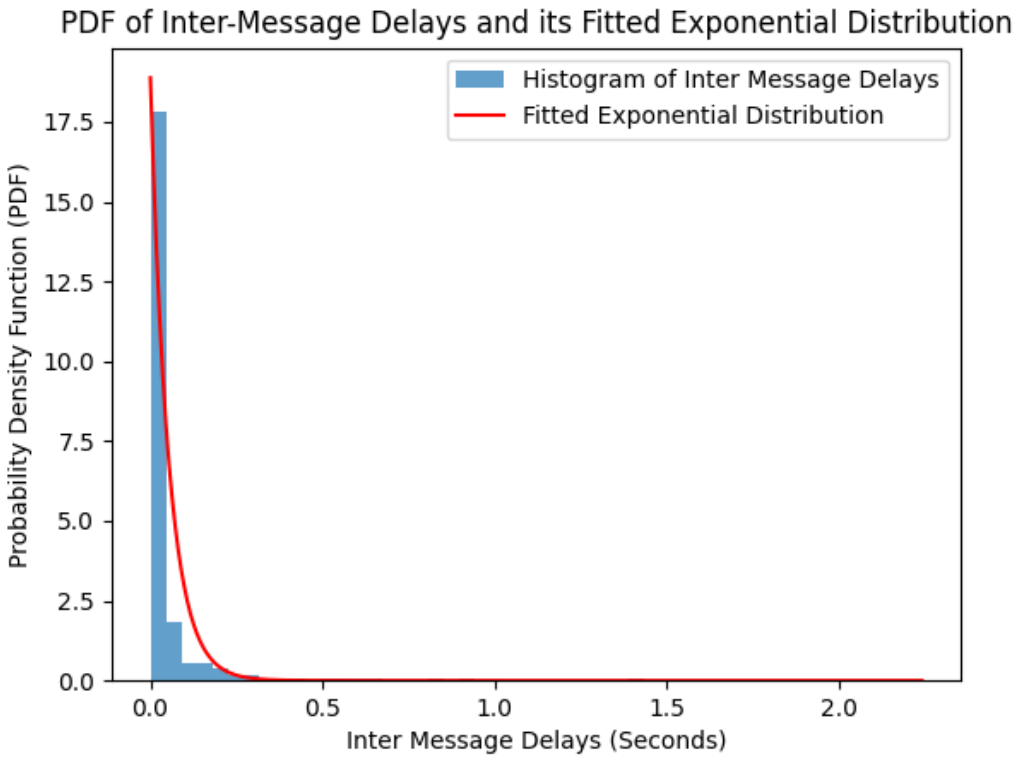
ההשהיות בין ההודעות בכל התרשימים מציגות התפלגות דומה בצורה משמעותית.
זה מצביע על כך שהדינמיקה של שליחת ההודעות היא דומה בין הקבוצות, לפחות מבחינת ההשהיות בין ההודעות.
הגודל הכולל של ההודעות בכל התרשימים דומה גם כן. זה מצביע על כך שההודעות בכל הקבוצות הן בגדלים דומים, כאשר כל סוג הודעה (אודיו, תמונה, וכו') מגיע לגודלו הטבעי.

ניתוח:

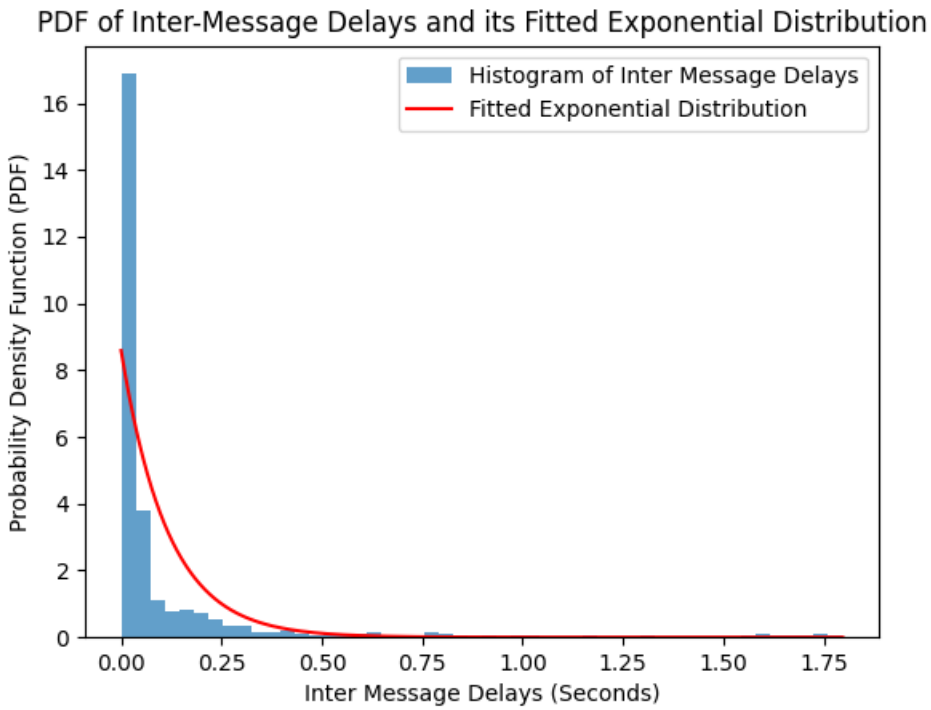
התרשימים החדשים מציגים את ההתפלגות של ההשהיות בין ההודעות וגם את התפלגות הגדלים של ההודעות בצורה שונה מהתרשימים הקודמים. ההבדל הוא בסוג התרשים אך המידע הכללי נשאר זהה.
המידע המוצג בתרשימים הללו יכול להקשות על מתקיף אם הוא מנסה להסיק מידע מהתרשימים, מאחר שההתפלגות מוצגת בצורה שונה יכולה לבלבל את המתקיף.

הקלטות עם רעשי רקע fig 3:

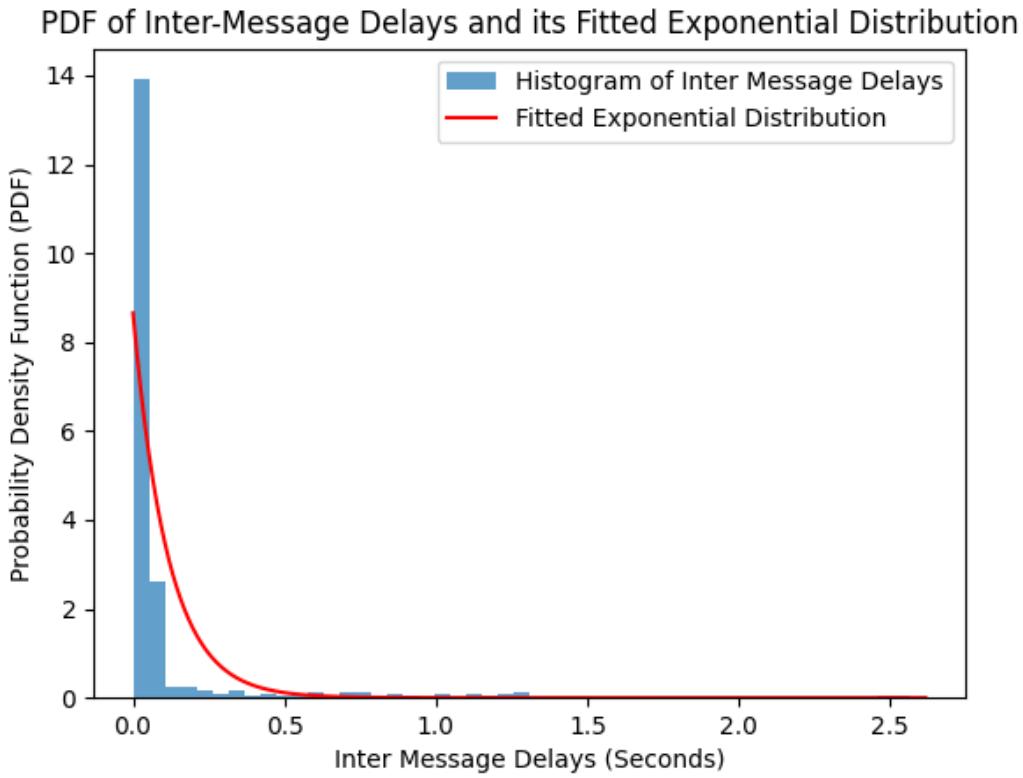
תמונות:



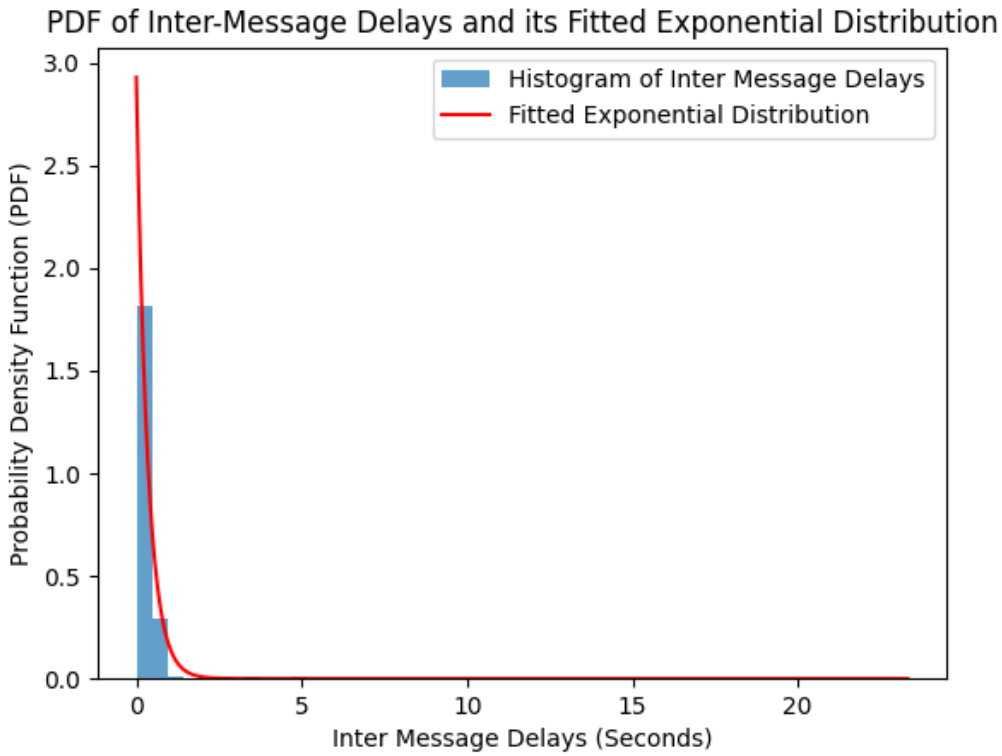
הקלטות:



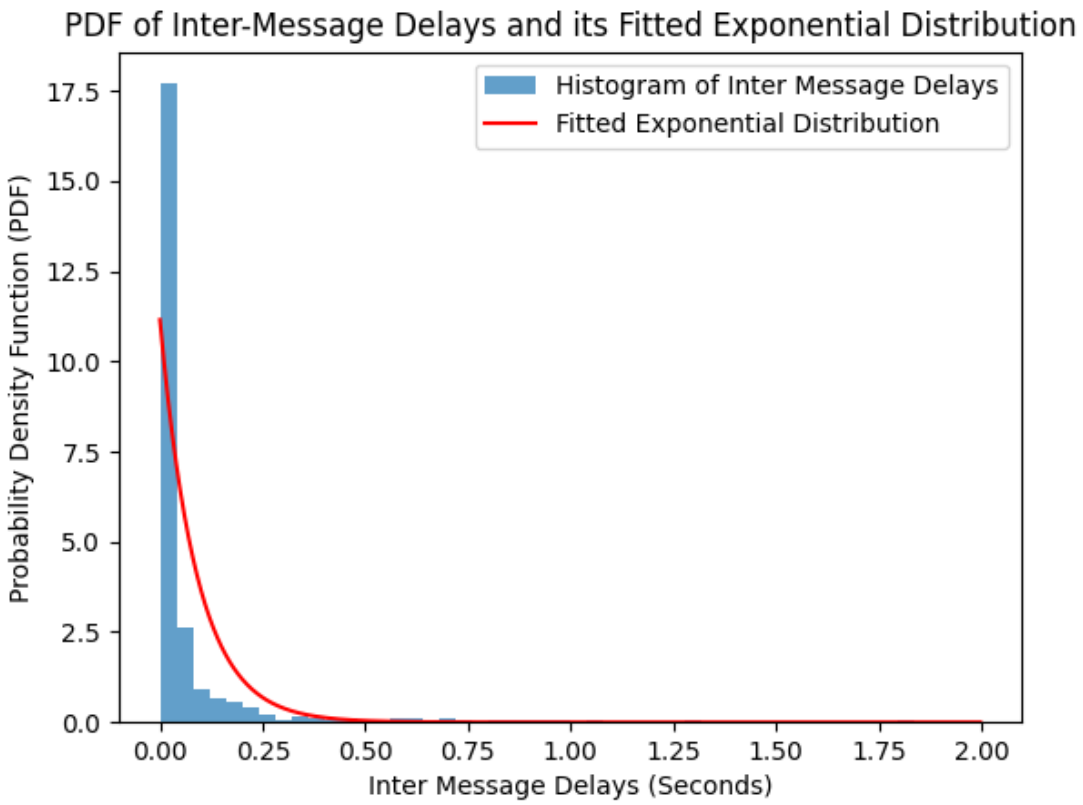
קבצים:


PDF of Inter-Message Delays and its Fitted Exponential Distribution
Histogram of Inter Message Delays
Fitted Exponential Distribution
Probability Density Function (PDF)
Inter Message Delays (Seconds)
0
2
4
6
8
10
12
14
0.0
0.5
1.0
1.5
2.0
2.5

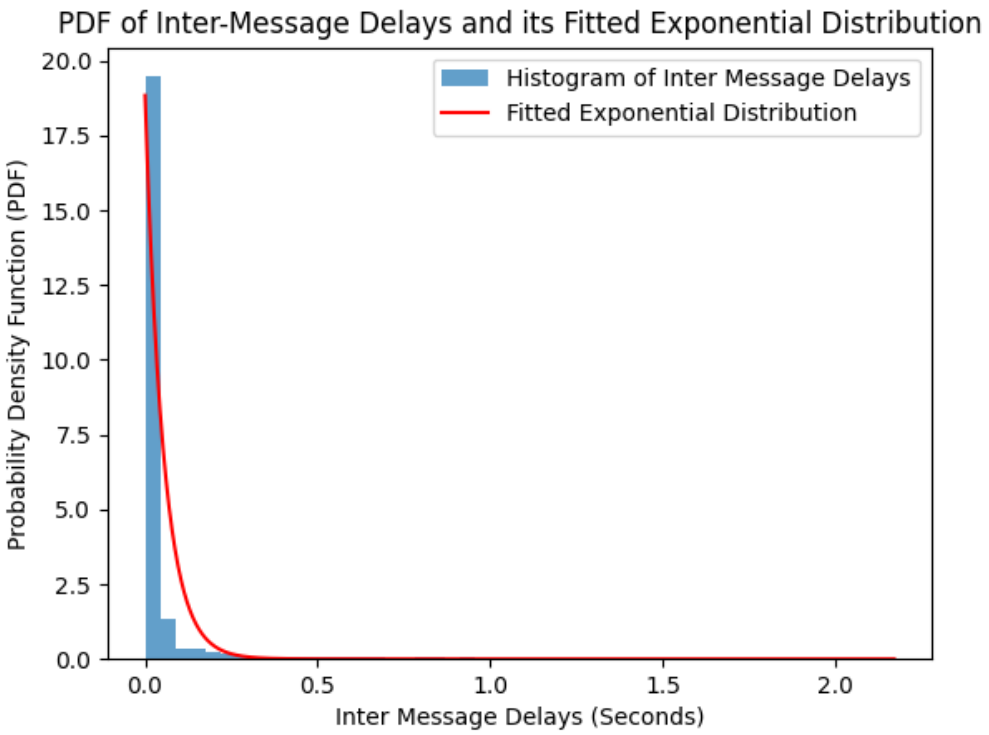
טקסט:


PDF of Inter-Message Delays and its Fitted Exponential Distribution
Histogram of Inter Message Delays
Fitted Exponential Distribution
Probability Density Function (PDF)
Inter Message Delays (Seconds)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
5
10
15
20

וידיאו:



הכל הכול:



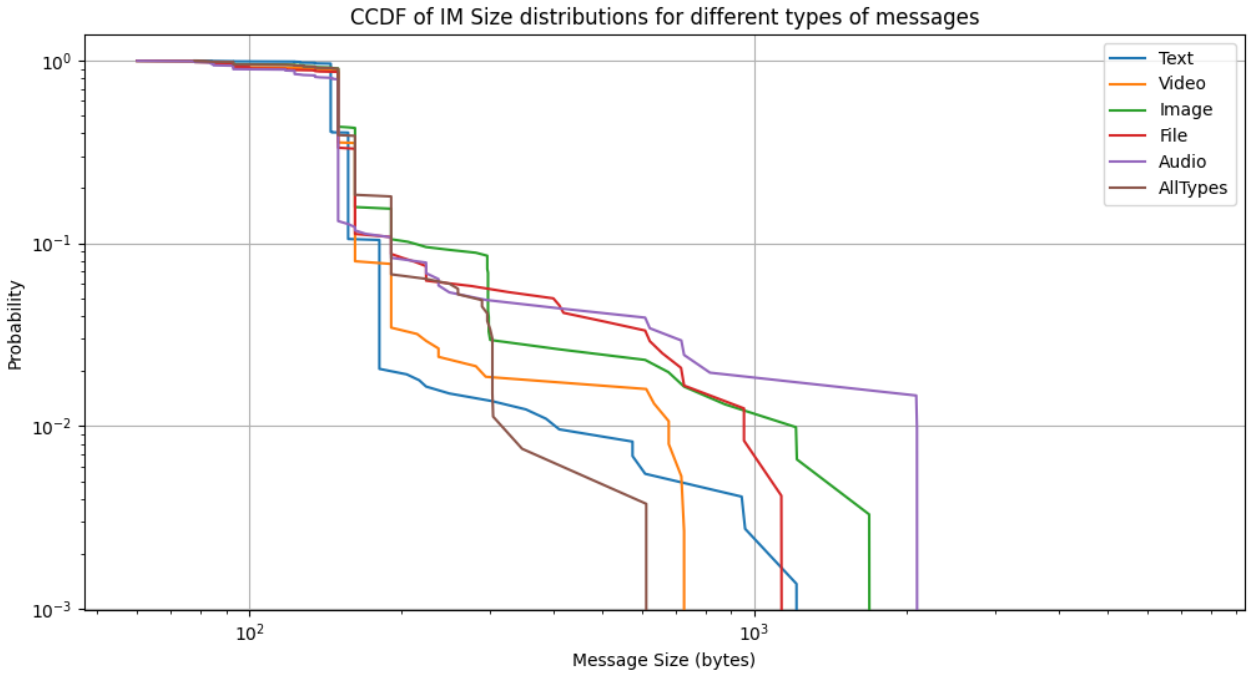
הסבר על הגרפים:

all_noise3.png (כל ההודעות ביחד עם רעשי רקע)
ההתפלגות של השהיות בין ההודעות דומה לתרשימים הקודמים.
הגודל הכולל של ההודעות דומה גם כן לתרשימים הקודמים, אך יש רעש נוסף שמופיע בתרשים.

aud_noise3.png, file_noise3.png, img_noise3.png, txt_noise3.png, vid_noise3.png (הודעות אודיו, קבצים, תמונה, טקסט, וידאו עם רעשי רקע)
בכל התרשימים, ההתפלגות של השהיות בין ההודעות והגודל הכולל של ההודעות הן דומות לתרשימים הקודמים.
עם זאת, הרעש שמופיע בתרשימים גורם לכך שהמידע מתפלג בצורה שונה וישנם שיאים נוספים.

ניתוח:
התרשימים החדשים עם רעשי הרקע מצגים את ההתפלגות הדומה של השהיות בין ההודעות והגדלים של ההודעות, אך הרעש המוסף גורם להתפלגות שונה בתרשימים.
הרעש יכול להקשות על מתקיף אם הוא מנסה להסיק מידע מהתרשימים, כאשר ההתפלגות המשתנהת והשיאים הנוספים יכולים להטעות או להבלבל את המתקיף.

כל הקבוצות באותו גרף fig 4:



הסבר:

תרשים ccdf4.png:

התרשים מציג התפלגות CCDF (הגרף מתאר את ההסתברות כתלות בגודל ההודעה).
לגבי השהיות בין ההודעות. נראה כי ישנם שני סוגים של התפלגויות בתרשים, כאשר כל אחת מהן מצביעה על דפוס שונה של שליחת הודעות:

התפלגות אחת (נראית כמו הקו הכחול): מצביעה על השהיות בינוניות בין ההודעות, שהיא דומה לאותם דפוסים שראינו בתרשימים הקודמים.
התפלגות שנייה (נראית כמו הקו האדום): מצביעה על השהיות ארוכות יותר בין ההודעות, מה שיכול להצביע על קבוצה אחרת או על דפוס שיחה שונה.

ניתוח:

התרשים מציג דפוס של שליחת הודעות בשני סגנונות שונים. זה יכול להצביע על כך שישנם שני סוגים של דיאלוגים או שיחות בקבוצה המוצגת בתרשים, או שזה מצביע על שני קבוצות שונות. ההבדל בדפוסים יכול להיות השפעה של הנושא של השיחה, ההשתתפות של המשתמשים, או אף מגבלות טכניות.

תשובות לשאלות:

האם ישנם מאפיינים ייחודיים לכל קבוצה?

כן, כל סוג של הודעה (אודיו, קובץ, תמונה, טקסט, וידאו) מצביע על מאפיינים מובחנים בגודל ההודעה בהשהיות בין ההודעות. לדוגמה, הודעות טקסט מתאפיינות בהחלפות תדירות ובגדלים קטנים, בעוד הודעות וידאו וקובץ הן בגדלים גדולים יותר ונשלחות באופן פחות תדיר.

האם אפשר להסיק אילו קבוצות אתה משתתף בהן באמצעות הטכניקות המפורטות במאמר?

במקרה שבו המשתמש המותקף פעיל תמיד בקבוצת הודעות מיידיות אחת בלבד:
אם המתקיף יכול לזהות את סוג ההודעות הנשלחות, הוא יכול להסיק את האופן בו המשתמש המתוקף משתתף בקבוצה. לדוגמה, קבוצה שבה נשלחות הרבה הודעות תמונה עשויה להיות קבוצת צילום.

במקרה שבו המשתמש המותקף פעיל במספר קבוצות הודעות מיידיות באופן יחד:
ההבחנה היא מורכבת יותר. כאשר ההודעות מגוונות מכמה קבוצות, יהיה קשה להסיק אילו הודעות מגיעות מאילו קבוצות, אלא אם המתקיף יכול לזהות דפוסים ייחודיים לכל קבוצה.

לסיכום, ישנם מאפיינים ייחודיים לכל קבוצה, אך היכולת להסיק אילו קבוצות המשתמש משתתף בהן תלויה בכמות המידע הזמינה למתקיף ובטכניקות שהוא משתמש בהן.