פרויקט גמר רשתות תקשורת - שירה צ'שלר 323825059,ואוהד וולפמן 316552496

חלק א (Dry Part) - סיכום מאמר

שם המאמר: Practical Traffic Analysis Attacks on Secure Messaging Applications

אפליקציות התכתבויות (IM=instant messaging) מאובטחות כמו טלגרם, סיגנל ו-ווטסאפ הפכו לחלק בלתי נפרד מהתקשורת היומיומית. במאמר הנ"ל דרשו החוקרים להעלות את המודעות לגבי האבטחה של האפליקציות המדוברות. החוקרים מסבירים שעל אף העובדה שאפליקציות אלו מוצפנות ולא ניתן לראות את תוכן ההתכתבויות, אפילו הצפנה מתקדמת, עשויה שלא להספיק כדי להגן באופן מלא על מידע משתמש רגיש מפני יריבים עם הכלים והטכניקות הנכונות אשר מטרתם לזהות (על ידי זיהוי כתובות IP) של חברים או מנהלים של קבוצת התכתבויות מסוימת.

החוקרים גילו כי כל פעולה שמתבצעת בתוך יישומים אלה - כמו שליחת הודעה, העלאת קובץ או אפילו רק הקלדה - מייצרת זרימת נתוני תקשורת עם דפוסים ייחודיים, והדגימו כי על ידי מעקב אחר התעבורה ברשת ניתן על ידי מודלים סטטיסטים לזהות את הדפוסים, ועל ידיהם לספק שפע של מידע רב ערך על משתמשים ופעילויותיהם. החוקרים מגדירים זאת כ'Traffic Analysis Attack'.

לתוקף ישנן מספר דרכים אופציונליות להשיג ground truth על תעבורת הערוץ כלומר להבין את דפוסי זרימת התעבורה של תוכנת מסרים מסוימת לפי האפשרויות של סוג קבוצת היעד:

א. אם קבוצת היעד הינה ציבורית (כלומר פתוחה לכל מי שרוצה להצטרף), התוקף יכול פשוט להצטרף לקבוצה כחבר. ברגע שהוא מצטרף, מעבר ליכולת שלו להסתכל בזמן אמת, הוא יכול להקליט את ההודעות שנשלחו יחד עם הmeta data שלהן (כגון זמן וגודל ההודעות), ולהשתמש במידע זה. סוג זה של נתונים יכול לתת ליריב תובנות לגבי מתי הקבוצה הכי פעילה, כמה חברים משתתפים ואולי אפילו באילו סוגי נושאים דנים (בהתבסס על גודל ותדירות ההודעות).

ב. אם קבוצת היעד אינה ציבורית אך התוקף כן הצליח להצטרף אליה והוא יכול לשלוח הודעות - הוא יכול לפרסם הודעות משלו עם דפוסי תעבורה ברורים (הידועים לו מראש). על ידי ניתוח האופן שבו חברים בערוץ מגיבים להודעות אלו, היריב יכול לקבל מידע נוסף על תנועת הערוץ.

ג. אם קבוצת היעד הינה פרטית, והתוקף אינו מצליח להצטרף אליה, אך הוא הצליח לזהות את כתובת הIP של אחד המשתתפים/ המנהלים בה, הוא יכול להאזין לתעבורת הרשת של המשתתף המזוהה. נתונים אלה יאפשרו לו לתעד את דפוסי התעבורה של המשתתף, מה שיכול לספק רמזים לגבי הפעילות בערוץ, גם אם אין ביכולתו לראות את ההודעות שנשלחות בפועל.

ביכולתו של התוקף לבצע האזנת סתר בתעבורת הרשת של המשתתף המזוהה במספר דרכים על ידי:

א. האזנה לתעבורת הרשת של ספק האינטרנט (Internet Service Provider - ISP), אשר דרך הנתבים שלו עוברת כל התעבורה.

ב. האזנה ל IXP (Internet eXchange Point) השרתים שחברות ISP מתחברות אליהן ולשרתי CDN (Content Delivery Network) המגבירים את מהירות הגעת מידע הרשת למשתמש.

ג. האזנה למשתמשים ספציפיים ע"י השגת צו האזנה נסתרת, למשל במידה והתוקף/ יריב הינו חברה ממשלתית ולכן יכול להשיג היתר משפטי שכזה.

החוקרים השתמשו בהתקפה של מעין Flow correlation - ניסיון לקשר בין זרימה ברשת לבין מאפייני התעבורה (זמנים של packets וגודליהן). בטבלה הנ"ל מוצגים נתונים וסטטיסטיקות של 5 סוגי מסרים:

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

לאחר המידול, החוקרים מציעים שני אלגוריתמי גילוי לפיהם ניתן לעקוב אחר תעבורה של קבוצה ולהחליט האם משתמש מסוים אכן חבר בקבוצה:

א. האלגוריתם הראשון הוא 'Event-Based Detector' (מזהה מבוסס אירועים).
Event (אירוע) יקרה כאשר תבוצע שליחה שתגרום לפרץ של packets בגודל מקסימלי - MTU (Maximum Transmission Unit, גודל המנה המקסימלי ששכבה נתונה בפרוטוקול מסוים יכולה להעביר). ואז בהינתן שלתוקף ישנה גישה לתעבורת הקבוצה הרצויה, ותעבורות משתמשים נוספים, באמצעות פונקציית קורלציה – ניתן לבדוק האם משתמש שייך לקבוצה או לא (אם הקורלציה בין התעבורה שלו בעלת קורלציה גבוהה מספיק, מעל threshold מסוים, לזו של הקבוצה לה מאזין התוקף).

ב. האלגוריתם השני הוא Shape-based Detector (אלגוריתם מבוסס-צורה), המתאם את צורות זרימת תעבורת ה-SIM על מנת לשייך את המשתמשים לערוצי היעד. 'צורה של תעבורה', לפי החוקרים, הכוונה לוקטור של אורך הפקטות כתלות בזמן. אלגוריתם זה הוא איטי יותר אך מציע ביצועי זיהוי מדויקים יותר מאלגוריתם מבוסס אירועים (מספיקות לו רק 15 דקות של תעבורת טלגרם כדי לזהות את המנהל של ערוץ SIM מסוים עם דיוק של 94%), והוא ומכיל 4 שלבים:

1) חילוץ אירוע, כמו באלגוריתם הראשון – על ידי זיהוי של פרצים של השימוש ברוחב הפס.
2) נרמול צורות התעבורה על ידי החלפת כל אירוע (כלומר, כל פרץ) בפס תנועה שרוחבו הוא te 2, כאשר te הוא הסף המשמש במהלך חילוץ אירוע. הנרמול מסיר את ההשפעה של רוחב הפס של המשתמש. גובה כל bar בגרף נבחר כך שהשטח מתחת לbar יהיה שווה לגודל האירוע. לבסוף אנו מחלקים כל bar לbins קטנים יותר ברוחב הts, ועם גובה השווה לגובה הפס המקביל של הקבוצה הנבדקת. לכן כל bar מורכב ממספר bins בעלי רוחב וגובה שווים. צורת התנועה החדשה תהיה הווקטור של גבהים של bins לאורך זמן.
3) בדיקת קורלציה של צורות התעבורה המנורמלות לזו של הקבוצה הנבדקת.
4) השוואה ביחס לtreshold מוגדר שתביא לחיזוי ביחס לכתובת המודגמת.

(נעיר כי בפועל, התוקף יכול לשלב את שני האלגוריתמים כדי לאזן את הtrade-off בין ייעול עלות החישוב (והמדרגיות) לעומת ביצועי (אחוז) הזיהוי)

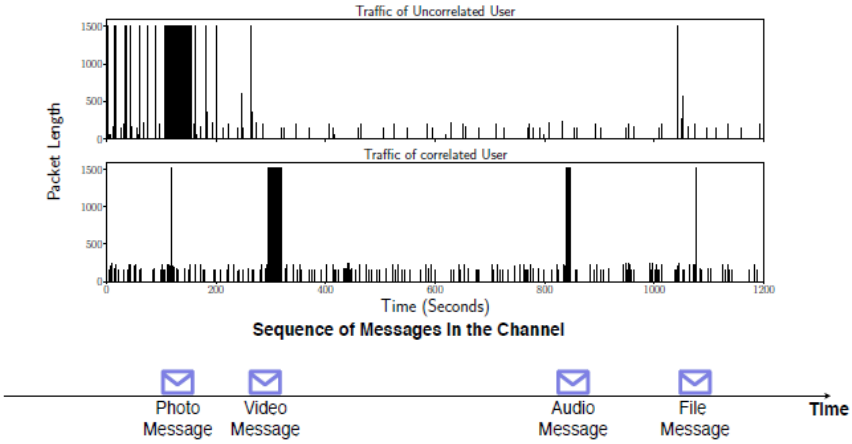

Fig. 8: Event extraction: IM Messages sent/received by a target user create bursts of (encrypted) packets; the adversary can extract events from packet bursts.

באיור 8 במאמר ניתן לראות תרשימים של גודל הpacket בקרב שני משתמשים – אחד שלא שייך לקבוצה (שעל כן הקורלציה בינו ובין התעבורה בקבוצה אינו גבוה, וגרף אחד של משתמש שאכן שייך לקבוצה ובעל קורלציה גבוהה לתעבורה של הקבוצה אל מול ציר זמן.

בתחתית האיור ניתן לראות events שקורים במהלך פרק הזמן בו מתבוננים על תעבורת שני המשתמשים (שליחת תמונה, סרטון, קובץ אודיו, וקובץ), כאשר עבור המשתמש אשר זוהה כשייך לקבוצה – ניתן לראות שבכל שליחה שכזו יש מספר MTU (לפחות 1, בדיוק בזמן הevent), בעוד שעבור המשתמש שלא זוהה כשייך לקבוצה ישנם MTU בזמנים שונים, לא בהכרח בזמן event.

לאחר מכן, בודקים החוקרים את המודלים שלהם על טלגרם, ווטסאפ וסיגנל. הם מראים שהפעלת האלגוריתמים שלהם נותנים תוצאות זיהוי טובות. דבר זה מהווה איום משמעותי על המשתמשים, לאור

הניסיונות ההולכים וגדלים של ממשלות מדכאות לפצח את הערוצים השנויים במחלוקת בפלטפורמות הללו. המחקר מהווה קריאת השכמה הן למשתמשים והן לספקים של שירותי הודעות כאלה. המשתמשים צריכים להיות מודעים לסיכונים ולהתאים את דפוסי השימוש שלהם בהתאם. במקביל, זה גם מדגיש את הצורך של ספקי שירותים לשלב אמצעי נגד אפקטיביים של ערפול תעבורה במערכות שלהם, מעבר להצפנה, כדי להבטיח את הפרטיות והבטיחות של המשתמשים שלהם.

לבסוף, המחברים מציגים מערכת אמצעי נגד הנקרא IMProxy שיכולה לשמש לקוחות IM ללא צורך בתמיכה כלשהי, אשר מביאה לתוצאות טובות אשר הוכחו גם בניסויים. מערכת זו נועדה להגן מפני ההתקפות מסוג זה בדיוק. החוקרים גילו כי ביצוע של tunnelling (מנהור) של תעבורת SIM דרך VPN וערבוב שלה עם תעבורת גלישה באינטרנט מפחיתה את דיוק ההתקפה באמצעות שני האלגוריתמים שלהם מ-93% ל-70%, והוספת תעבורת כיסוי (cover traffic) עם תקורה של 17% מורידה את הדיוק ל-62%.

<u>קישורים</u>

לינקד אין –

1. https://www.linkedin.com/in/shira-chesler-4438b5222/
2. https://www.linkedin.com/in/ohad-wolfman/

גיטהאב –

https://github.com/ohadwolfman/Networks_Final_Project