דו"ח פתיחה
פרויקט באבטחת מידע 236349
סמסטר חורף 2018-2019

נושא הפרויקט: Acoustic keylogger

מגישים:

- יוסי גורשומוב, 308561562, yossigor@campus.technion.ac.il.
- משי פריד, 313574972, meshifried@campus.technion.ac.il.

מנחים:

- עמיר שוורץ, amir.schwartz@citi.com.
- דני טילמן, danny.tylman@citi.com.

תיאור הפרויקט:

בפרויקט זה נחקור התקפת ערוץ צד (side channel attack) על האותות האקוסטיים של הקשות מקלדת ונממש אפליקציה על רכיב מחשוב נייד (android smartphone או raspberry pi) המממשת את ההתקפה שמטרתה חשיפת המידע שהוקלד ע"י הקורבן ובפרט חשיפת רצפים אקראיים כגון סיסמאות.

מטרת הפרויקט:

מטרת הפרויקט היא בניית אב טיפוס של Acoustic keylogger, והפקת דו"ח המפרט את איכות תוצאות ההתקפה שבוצעה באמצעותו, כלומר את איכות זיהוי המידע המוקלד מתוך האות האקוסטי של ההקלדות.

הנושאים שיטופלו במהלך הפרויקט:

1. נחקור דרכים לחילוץ תכונות (feature extraction) מתוך האות האקוסטי של הקשות מקלדת.
2. נבחן אלגוריתמי למידה שמטרתם לבנות מסווג עבור ההקשות ובפרט נבחן את האפשרות לביצוע unsupervised learning שיאפשרו את ביצוע הלמידה עבור הקלטות שאינן מסומנות מראש (אקט שדומה באופיו ל ciphertext only attack).
3. נבצע ארכיטקטורה לאב טיפוס עבור אפליקציה המממשת את ההתקפה.

4. נממש מספר אלגוריתם נבחרים על גבי האפליקציה במכשיר נייד.
5. נבחן מדדים עבור הצלחה של התקפות באמצעות האפליקציה. בפרט, נבחן מהו אחוז הזיהוי עבור הקלדות של טקסטים בשפה האנגלית, ומהו אחוז הזיהוי עבור רצפים אקראיים כמו סיסמאות.
6. בהתאם לתוצאות ההתקפה נבחן כיצד ניתן לשפר את הצלחת ההתקפה ומה הגורמים המשפיעים על אחוז הזיהוי (הכוונה שנבחן פרמטרים פיזיים כמו למשל מקלדות שקטות, דפוסי הקלדה שונים, איכות אמצעי ההקלטה וכדומה ובנוסף פרמטרים שאינם פיזיים כמו אורך סיסמאות והתוכן שלהם).
7. נעלה הצעות כיצד ניתן להתגונן מפני התקפות מן הסוג הזה.

תיאור כללי של ההתקפה:

ההתקפה תכלול שני שלבים עיקריים:

1. שלב האימון:
   a. התוקף יקליט את הקשות המקלדת של הקורבן ע"י האפליקציה.
   b. האפליקציה תבצע תהליך של feature extraction שלאחריו כל הקשה תיוצג על ידי וקטור תכונות.
   c. האפליקציה תבצע תהליך של unsupervised learning ותבנה מסווג עבור הקשות המקלדת על בסיס מודל עבור השפה האנגלית.
2. שלב הזיהוי:
   a. התוקף יקליט את הקשות המקלדת של הקורבן ע"י האפליקציה, רצף ההקלדות הזו הוא הרצף שהתוקף רוצה לזהות את תוכנו (למשל סיסמא או מידע מסווג).
   b. האפליקציה תבצע תהליך של feature extraction שלאחריו כל הקשה תיוצג על ידי וקטור תכונות.
   c. האפליקציה תכניס את וקטורי התכונות של ההקשות למסווג ותזהה אותן על בסיס מודל עבור השפה האנגלית.
   d. - *אופציונלי* - ההקשות המסומנות בהסתברות גבוהה יוכנסו כfeedback למסווג על מנת לשפר את שלבי הזיהוי הבאים.

מאמרים הקשורים לנושא:

1. Keyboard acoustic emanations D Asonov, R Agrawal - 2004 computer.org , Cited by 278 , PDF
2. Keyboard acoustic emanations revisited L Zhuang, F Zhou, JD Tygar - ACM Transactions on Information and …, 2009 - dl.acm.org , Cited by 279 , PDF
3. Dictionary attacks using keyboard acoustic emanations Y Berger, A Wool, A Yeredor - Proceedings of the 13th ACM conference …, 2006 - dl.acm.org, Cited by 121 , PDF

תוצאות צפויות:

- על פי התוצאות שמוצגות במאמר של L Zhuang, F Zhou (מאמר מס' 2 ב*מאמרים הקשורים לנושא*) אנו צופים שבניית Acoustic keylogger המבצע את ההתקפה המתוארת באחוזי הצלחה סבירים היא אפשרית בהחלט בייחוד עבור זיהוי של רצפי הקלדות בשפה האנגלית, **הציפייה שלנו היא לבנות אב טיפוס על פי המודל שמוצג במאמר הנ"ל המסוגל לזהות <u>80% מהקשות</u> המקלדת של קורבן המקליד בשפה האנגלית כאשר הקלט לאלגוריתם הלמידה הוא הקלטת אודיו לא מסומנת באורך <u>10 דקות בלבד</u> המכילה רצפי הקשות קודמים של הקורבן.**
- אנו מצפים שהשפעתם של שינויים פיזיים כמו אנשים שונים ומקלדות שונות ישפיעו על איכות ההתקפה באופן משמעותי. כלומר, לאחר שנבנה מסווג בהנתן הקלטות של אדם A ומקלדת B:
  - סיכוי הזיהוי של הקשות אדם A ומקלדת שונה מB תהיה נמוכה משמעותית מ80%.
  - סיכוי הזיהוי של של הקשות אדם שונה מA על מקלדת B תהיה נמוכה משמעותית מ80%.
- אנו מצפים שסיכויי ההצלחה בזיהוי של רצפים אקראיים כמו סיסמאות הם נמוכים משמעותית מזיהוי הקשות של רצפים בעלי משמעות בשפה האנגלית.

לוח זמנים:

לוח הזמנים של הפרויקט מחולק לשלבים הבאים (ומתאים לסעיפים המוצגים *בנושאים שיטופלו במהלך הפרויקט*):

1. <u>כתיבת דו"ח פתיחה (שבוע ראשון של הסמסטר 21-27/10)</u>
2. <u>מחקר וסקר ספרות ותוצאות קודמות (שלושה שבועות 28/10-17/15):</u>
   a. משימה זו מתאימה לסעיפים 1 ו 2 *בנושאים שיטופלו במהלך הפרויקט*.
   b. בשלושת השבועות הללו נבצע סקירה של החומר הרלוונטי עבור הפרויקט ובפרט נבחן את שלושת המאמרים שהוצגו במסמך זה.
   c. בעקבות משימה זו נבחר אלגוריתמים בעלי ביצועים מיטביים עבור הבעיה הנתונה שאותם נממש במהלך הפרויקט.
3. <u>תכנון ארכיטקטורה עבור אב טיפוס של Acoustic keylogger (שבועיים 18/11-1/12):</u>
   a. משימה זו מתאימה לסעיף 3 *בנושאים שיטופלו במהלך הפרויקט*.
   b. בשבוע זה נבצע ארכיטקטורת תוכנה עבור אפליקציית הAcoustic keylogger.
   c. נבחר את הסביבה שבה נרצה לממש את אב הטיפוס ומהן שפות התכנות והסיפריות הרלוונטיות למשימה זו.
   d. נכתוב מסמך המפרט את ארכיטקטורת התוכנה.
   e. הצגת הארכיטקטורה למנחים וקבלת משוב.
4. <u>מימוש האפליקציה וכתיבת אב טיפוס - שלב א' (שלושה שבועות 2/12-22/12):</u>
   a. משימה זו מתאימה לסעיף 4 *בנושאים שיטופלו במהלך הפרויקט*.
   b. הוכחת יכולת של feature extraction וכתיבת המודול שמבצע את הפעולה הנ"ל.
   c. הוכחת יכולת למידה ובניית מסווג.
   d. כתיבת בדיקות יחידה וסקרי קוד.
5. <u>ביצוע מדדים עבור המימוש בשלב א' (שבוע אחד 23/12-29/12):</u>
   a. משימה זו מתאימה לסעיף 5 *בנושאים שיטופלו במהלך הפרויקט*.
   b. בדיקת איכות הסיווג של התוצר של שלב א' בפיתוח.
   c. בהנתן והתוצאות אינן סבירות יש לבחון כיצד ניתן לתת להן פתרון.

d. הצגת התוצאות למנחים וקבלת משוב על איכות התוצאות ושינויים נדרשים.

6. מימוש האפליקציה וכתיבת אב טיפוס - שלב ב' (שלושה שבועות 2/12-22/12):
   a. משימה זו מתאימה לסעיף 4 *בנושאים שיטופלו במהלך הפרויקט*.
   b. טיפול בהערות ובשינויים הנדרשים שעלו בבחינת המדדים.
   c. העברת המימוש לסביבה שיכולה לרוץ במכשיר נייד. כלומר, כתיבת אפליקציה למכשיר נייד המפעילה את תוצרי שלב א'.
7. ביצוע מדדים עבור המימוש בשלב ב' (שבועיים 30/12-12/1):
   a. משימה זו מתאימה לסעיף 5 *בנושאים שיטופלו במהלך הפרויקט*.
   b. בדיקת איכות הסיווג של התוצר של שלב ב' בפיתוח.
   c. ביצוע סימולציה מלאה של התקפה (DEMO).
   d. בהנתן והתוצאות אינן סבירות יש לבחון כיצד ניתן לתת להן פתרון.
   e. הצגת התוצאות למנחים וקבלת משוב על איכות התוצאות ושינויים נדרשים.
8. כתיבת מצגת אמצע ופוסטר (שבוע אחד 13-19/1)
9. הצגת פוסטר (על פי אתר הקורס 22/1)
10. מימוש האפליקציה וכתיבת אב טיפוס - שלב ג' (שבוע אחד 20/1-26/1):
    a. משימה זו מתאימה לסעיף 4 *בנושאים שיטופלו במהלך הפרויקט*.
    b. טיפול בהערות ובשינויים הנדרשים שעלו בבחינת המדדים.
    c. שיפור הקוד.
    d. שיפור ממשק המשתמש.
    e. בחינת שימוש בחישוב cloud עבור תהליך הלמידה.
    f. בחינת מימוש סעיפים שהוגדרו אופציונליים במהלך הפרויקט.
11. מועדי א' (תאריכים 26/1-14/2)
12. מועדי ב' (תאריכים 28/2-14/3)
13. כתיבת דו"ח סיום (תאריכים 15/3-22/3)
    a. ניתוח התוצאות על פי סעיפים 5,6 ו 7 ב*בנושאים שיטופלו במהלך הפרויקט*.
    b. שליחת הדו"ח ואב הטיפוס למנחים לקבלת משוב.
14. הגשת הפרויקט (תאריכים 1/4)

נושאים נוספים:

1. מילואים:
   a. יוסי יהיה במילואים במהלך החודש הראשון של הסמסטר (22/10-15/11) ולכן לא יהיה במפגש הפתיחה ויוכל להגדיש זמן מוגבל עבור העבודה על הפרויקט.
2. שיתוף מידע וניהול תצורה:
   a. מסמכי הפרויקט יכתבו בעברית בפלטפורמת google docs.
   b. תהליך הפיתוח מנוהל בפרויקט פרטי בgithub בלינק הבא: https://github.com/yossigor/AcousticKeylogger .
   c. יקבעו שיחות מעקב באופן שוטף עם המנחים.
3. סביבת עבודה ומשאבים נדרשים:
   a. עבודת המחקר והפיתוח תתבצע ברובה על בסיס המחשבים הפרטיים של הסטודנטים:
      i. ניסויים ראשוניים יתבצעו על מחשב dell inspiron 17r 5737.

ii. ניסויים במכשיר נייד יהיו על בסיס Android smartphone מדגם nexus 5 (קיימים לנו 2 מכשירים מדגם זה).
iii. הקלטות של אותות קול יתבצעו ע"י מיקרופון שקיים במכשירים הנ"ל.

b. דרישות נוספות שיתכנו במהלך הפרויקט ולא קיים לנו מענה אליהם:
   i. מכשיר נייד מסוג Raspberry pi.
   ii. חשבון עבור שירות מחשוב ענן.