פרויקט באבטחת מידע - Acoustic Keylogger

סקר ספרות ומחקרים קודמים

מאמרים שנסקרו

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) Cracking Passwords using Keyboard Acoustics and Language Modeling A Kelly - University of Edinburgh 2010, PDF

Keyboard acoustic emanations

Keyboard acoustic emanations

במאמר זה מובאת התקפת ערוץ צד על האות האקוסטי של הקשות המקלדת.

דרישות ההתקפה:

- **הקלטה מתויגת של הקשות המקלדת של הקורבן** (כלומר, צמדים של אות הקשה והתו המוקש). מהקלטה זו מופק training set עבור אלגוריתם הלמידה.
- **הקלטה שאינה מתויגת של הקשות המקלדת של הקורבן.** את הקלטה זו רוצים להזין למסווג ולקבל את המידע שהוקלד בה.

Keyboard acoustic emanations

שלב א', עיבוד האות האקוסטי:

- במאמר זה עיבוד האות האקוסטי בוצע ידנית, כלומר זיהוי מיקומו של הpeak שנוצר בעקבות ההקשה בוצע ידנית ע"י החוקרים.
- מחלונית הזמן שבה קיים הpeak יחולצו התכונות בשלב הבא.

Keyboard acoustic emanations

שלב ב', חילוץ התכונות:

- על חלונית הזמן שבה קיימת הקשה מבצעים FFT לקבלת וקטור התדרים שנשמעו בהקשה.
- וקטור המקדמים המנורמל הוא וקטור התכונות שיוזן לאלגוריתם הלמידה.
- נשים לב שבהתקפה המוצגת במאמר זה עבור כל וקטור מצומד התו שהוקש.
- בסוף שלב זה מתקבל training set מתוייג.

Keyboard acoustic emanations

שלב ג', למידה:

- בשלב זה training set הוזן לרשת נוירונים.
- לאחר האימון הרשת מוכנה לקבל וקטורי תכונות של הקשות חדשות.

Keyboard acoustic emanations

שלב ד', זיהוי:

- בהנתן אות של הקשה חדשה שאינה מתוייגת נבצע עליה את עיבוד האות וחילוץ התכונות בדומה לשלבים א' - ג' .
- כעת מוזן לרשת הנוירונים המאומנת וקטור התכונות.
- הרשת פולטת את סיווג הוקטור, כלומר את הניחוש עבור התו שהוקש.

Keyboard acoustic emanations

במאמר מוצג אחוז זיהוי של 79%. כלומר בממוצע 79 תווים מתוך 100 מזוהים נכון ע"י רשת הנוירונים המאומנת.

Keyboard acoustic emanations revisited

Keyboard acoustic emanations revisited

מאמר זה מרחיב את העבודה שבוצעה במאמר [1] באופן הבא:

- זיהוי מיקום ההקשות באופן אוטומטי ע"י עיבוד מקדים של האות.
- שימוש באלגוריתם cepstrum במקום FFT לחילוץ התכונות.
- שימוש במסווג לינארי במקום רשת נוירונים.
- שימוש במודל תלוי שפה על מנת לבצע אימון של מסווג שלא מצריך קבוצת אימון מתויגת.
- שימוש בזיהוי המסווג כfeedback לאלגוריתם הלמידה על מנת לשפר הזיהויים הבאים.

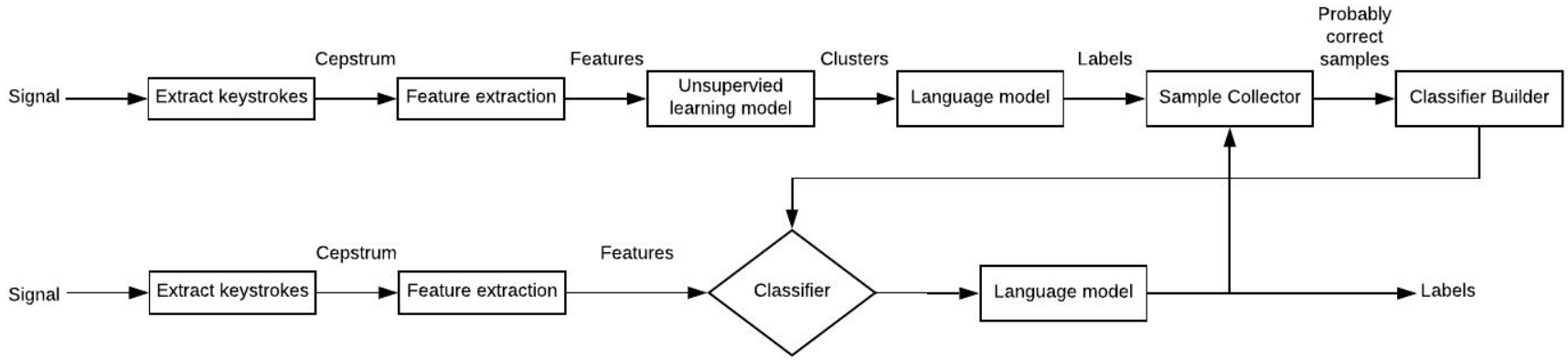
Keyboard acoustic emanations revisited

מאמר זה מקל את דרישות ההתקפה **ולא מצריך שימוש בהקלטה מתוייגת.**

<u>דרישות ההתקפה:</u>

- **הקלטה שאינה מתויגת של הקשות המקלדת של הקורבן באורך לפחות 10 דקות.** הקלטה זו תשמש אותנו בשלב הלמידה.
- **הקלטה נוספת שאינה מתויגת של הקשות המקלדת של הקורבן.** את הקלטה זו רוצים להזין למסווג ולקבל את המידע שהוקלד בה.

Keyboard acoustic emanations revisited



Keyboard acoustic emanations revisited

<u>עיבוד האות האקוסטי באופן אוטומטי:</u>

- במאמר מוצע אלגוריתם לזיהוי הpeak של ההקשה הרלוונטי עבור אלגוריתם הלמידה.
  - האלגוריתם מבצע וריאנט של STFT על מנת לחשב סכום של עוצמות התדרים בזמן.
  - סכום זה מסומן כ energy.
  - בהינתן והenergy עובר סף מסוים הנקבע ידנית מניחים כי בזמן זה הוקש תו.

Keyboard acoustic emanations revisited

אלגוריתם חלופי לחילוץ תכונות:

- במאמר מוצג אלגוריתם חלופי לFFT בשם Cepstrum. כותבי המאמר מציינים כי עבור בעיית הלמידה הספציפית הזו האלגוריתם החלופי נותן תוצאות טובות יותר.

Keyboard acoustic emanations revisited

מסווג לינארי אל מול רשת נוירונים:

- כותבי המאמר מציינים כי עבור הבעיה הנתונה מסווג לינארי מספק ביצועים טובים יותר מאשר רשת נוירונים.

Keyboard acoustic emanations revisited

מודל תלוי שפה:

- במאמר מצוין כי ניתן למעשה לחלק את הClusters שהתקבלו לתווים על פי שכיחות התווים בשפה האנגלית.
- למרות זאת במאמר מוצג מודל מורכב יותר המשתמש בHidden Markov Models על מנת לתייג את הClusters.
- לאחר תיוג הClusters ניתן לתת סיווג לכל הקשה ולקבל טקסט. את הטקסט מכניסים לאלגוריתם תיקון שגיאות. הפלט של אלגוריתם תיקון השגיאות נלקח בחשבון ובעזרתו מקבלים טקסט בעל נכונות טובה יותר.
- הטקסט המתקבל לאחר תיקון השגיאות מייצג training set מתוייג. את הtraining set הזה מכניסים למסווג הלינארי.

Keyboard acoustic emanations revisited

במאמר מוצג אחוז זיהוי של 96% עבור קלטים בשפה האנגלית.
בנוסף מסווג מאומן היטב יכול לזהות רצפים רנדומליים כמו סיסמאות בהסתברות טובה.

Cracking Passwords using Keyboard Acoustics and Language Modeling

Cracking Passwords using Keyboard Acoustics and Language Modeling

מאמר זה מציג את האלגורימים שהוצעו במאמרים [1] ו [2] באופן מפורט יותר. לכן אנחנו נתמקד בו כמקור המידע העיקרי עבור המימוש שלנו.