Informationssicherheit und IT-Forensik

– Übung –

UEB-04 – Entropie

Grundlagen

- einfach gesagt: Wir wollen messen, wie zufällig ein Passwort eigentlich ist
- Dazu wollen wir bestimmen, wie stark ein einzelnes Zeichen diesen Zufall erhöht
- Dieses Maß für den Zufall bzw. die Komplexität eines Passworts nennen wir „Entropie“

Definition

- Claude E. Shannon (der selbe Shannon wie aus der Vorlesungseinheit Kryptographie)
- „A Mathematical Theory of Communication“ (1948)
- Überlegungen dazu, wie man Informationen möglichst effizient kodieren und auf einem Datenkanal übertragen kann
- Definiert damit den Informationsgehalt von Zeichen

Definition: Informationsgehalt

- Informationsgehalt eines Zeichens x: I(x)
- Wie viele Bit braucht man, um ein Zeichen x zu übertragen?
- Man benötigt die Auftrittswahrscheinlichkeit P(x) eines Zeichens
- Um x dann binär darzustellen, benötigt man

$$I(x) = \log_2(1/P(x)) \text{ [bit]}$$

Definition: Informationsgehalt

$I(x) = \log_2(1/P(x))$ [bit]

- Beispiele:
  - Alice sendet zwei Zeichen, 1 und 0. Beide Zeichen treten gleich wahrscheinlich auf:
    - $P(1) = P(0) = 0.5$
    - $I(1) = I(0) = \log_2(1/0.5) = \log_2(2) = 1$ [bit]
  - Bob sendet zwei Wörter, „kalt“ und „heiß“. Beide Wörter treten gleich wahrscheinlich auf:
    -
    -

Definition: Informationsgehalt

$I(x) = \log_2(1/P(x))$ [bit]

- Beispiele:
  - Alice sendet vier Zeichen: „a“, „b“, „c“ und „d“. In 50% der Fälle sendet sie „a“, in 25% der Fälle „b“. „c“ und „d“ werden jeweils in 12,5% der Fälle versendet.
    - P(„a“) = 0.5, P(„b“) = 0.25, P(„c“) = P(„d“) = 0.125
    - I(„a“) = log2(1/0.5) = 1
    - I(„b“) = log2(1/0.25) = 2
    - I(„c“) = I(„d“) = log2(1/0.125) = 3

Definition: Entropie

Entropie H:

- Mittlerer Informationsgehalt eines Zeichens

Wir nehmen der Einfachheit halber an:

- Alle Zeichen eines Alphabets können gleichwahrscheinlich auftreten, dann ist H = I(x)
- Das gilt z.B. für Passwörter, die zufällig generiert werden
- Das gilt typischerweise nicht für Passwörter, die von Menschen gewählt werden

Entropie von Passwörtern

- Passwörter werden gewählt über ein festes Alphabet (etwa die Menge aller Kleinbuchstaben a...z)
- Es gibt 26^6 mögliche Passwörter mit 6 Stellen über diesem Alphabet
- Statt nun die Entropie einzelner Zeichen zu bestimmen, kann man auch die Entropie des gesamten Passworts berechnen:
  - $P(x) = 1/26^6$ – Auftrittswahrscheinlichkeit eines 6-stelligen Passworts, wenn man annimmt, dass jedes Passwort gleich häufig auftritt
  - $H = \log_2(1/P(x)) = \log_2(26^6) = 28{,}2$ [bit]

Entropie von Passwörtern

Beispiele

- 8-stellig, Groß- und Kleinschreibung: 52 Zeichen
  - $H = \log_2(52^8) = 45{,}6$ [bit]
- 8-stellig, mit Sonderzeichen und Zahlen: ~80 Zeichen
  - $H = \log_2(80^8) = 50{,}58$ [bit]
- 9-stellig, Groß- und Kleinschreibung: 52 Zeichen
  - $H = \log_2(52^9) = 51{,}30$ [bit]
- Nimmt man einen Wortschatz von ca. 65.000 Wörtern an, ergibt sich für ein beliebiges existierendes Wort (egal wie lang):
  - $H = \log2(65000) = 16{,}00$ bit

Eine einzige Stelle mehr bringt einen höheren Zuwachs als eine Erweiterung des Zeichenvorrats um 28 weitere Zeichen! Länge ist also viel wichtiger als Komplexität.

Entropie von Passwörtern

Ein einziges Zeichen mehr bringt einen höheren Zuwachs als eine Erweiterung des Zeichenvorrats um 28 weitere Zeichen! Länge ist also viel wichtiger als Komplexität.

- Auf dem Konzept basiert auch Diceware (correct horse battery staple)
- Kleine Liste von ~8000 kurzen Wörtern
- Ein einzelnes Wort hat ca. 12,9bit Entropie
- Die Kombination aus fünf Wörtern hat schon 64,6bit (5*12,9)
- 6: 77,5bit
- 7: 90,4bit

http://world.std.com/~reinhold/diceware.html

Entropie: Praxisbeispiele

Wie hoch ist die Entropie von…

- 20 Stellen, bestehend aus Kleinbuchstaben?

- 6 Wörter aus einer Liste von 2048 Wörtern?

- Wie viele Wörter aus dem Duden (27. Auflage) muss man aneinanderreihen, um mindestens 80bit Entropie zu erreichen?

Hausaufgabe bis Samstag (24.02.24)

- Fortentwicklung des Konzepts zur Einbruchmeldeanlage
- Bearbeiten Sie das Konzept gründlich, aber nicht zu ausführlich
  - zu jedem VL-Block ~15min Gedanken machen, wie das Gelernte in Ihrem Konzept berücksichtigt werden muss
  - Lassen Sie *auch* einfließen die Erkenntnisse aus Gastbeitrag Spyra/Gastbeitrag Hirschmeier:
    - Welche Herausforderungen und Hürden sehen Sie aus Sicht des Datenschutzes?
    - Wie viel Zeitaufwand planen Sie für die Dokumentation ein? Wo schätzen, sind aus Sicht der Security, aber auch der Safety, mögliche Fallstricke, die beim Erstellen einer Dokumentation entstehen können

      (Beispiel: Risiko, dass der Anwender die Dokumentation nicht versteht und die Alarmanlage am Ende eines Tages nur scheinbar einschaltet)
- Am Samstag wollen wir über das Thema gemeinsam diskutieren. Vorbereitung hilft dabei.
- Dazu möglicherweise auch eine Prüfungsfrage