Wolfgang K. Seiler
Tel. 2515

A5, C201

3. November 2004

2. Übungsblatt Kryptologie

Aufgabe 1: *(6 Punkte)*

Eine Quelle mit Alphabet $A = \{x_1, \ldots, x_n\}$ und Wahrscheinlichkeit p_i von $x_i \in A$ sende eine große Anzahl m von Zeichen aus. Dann kann man, ohne einen großen Fehler zu machen, davon ausgehen, daß das Zeichen x_i unter diesen m Zeichen $p_i m$-mal vorkommt und daß diese Zahl ganz ist. Somit können wir davon ausgehen, daß für jedes $x_i \in A$ bekannt ist, wie oft es vorkommt; unbekannt ist „nur“ die Reihenfolge der Zeichen. Zeigen Sie

a) Es gibt $\dfrac{m!}{(p_1 m)! \cdots (p_n m)!}$ verschiedene Möglichkeiten, diese Zeichen anzuordnen.

b) $\lim\limits_{m\to\infty} \dfrac{1}{m} \log_2 \dfrac{m!}{(p_1 m)! \cdots (p_n m)!}$ ist gleich der Entropie der Quelle.

(Hinweis: Nach der Stirling*schen Formel ist* $\log_2 N! = N \log_2 N - \frac{N}{\ln 2} + \frac{\log_2 N}{2} + O(1)$*, wobei* $O(1)$ *einen Term bezeichnet, der für* $N \to \infty$ *beschränkt bleibt.)*

Aufgabe 2: *(8 Punkte)*

Wählen Sie einen deutschen Klartext von mindestens 30 000 Buchstaben und berechnen Sie die Entropie und Redundanz pro Buchstabe auf Grund

a) des Wissens über Buchstabenhäufigkeiten

b) des Wissens über Kontakthäufigkeiten

bezüglich der folgenden Alphabete:

- Nur 26 Buchstaben
- ASCII-Zeichen

Welche Folgerungen ergeben sich daraus für die Kryptologie?

Aufgabe 3: *(6 Punkte)*

Führen Sie entsprechende Berechnungen wie in der vorigen Aufgabe (auf der Grundlage von ASCII-Zeichen) durch für

a) eine Datei mit deutschem Text in Word oder einem ähnlichen System

b) einer Programmdatei in Ihrer Lieblingsprogrammiersprache

c) einer ausführbaren Datei Ihres bevorzugten Betriebssystems

d) einer mit `zip` oder einem ähnlichen Programm komprimierten Datei!

Abgabe bis zum Dienstag, dem 9. November 2004, um 12.00 Uhr