Kryptologie

Die Autokoinzidenzindizes eines Textes

a7Hzq .#5r<
kÜ\as TâÆK$
ûj(Ö2 ñw%h:
Úk{4R f~`z8
¤˜Æ+Ô „&¢Dø

Definition

Für die Kryptoanalyse periodischer polyalphabetischer Chiffren ist die folgende Konstruktion besonders interessant:

Sei $a_{(q)}$ bzw. $a_{(-q)}$ der um q Stellen zyklisch nach rechts bzw. links verschobene Text, also

$$\begin{aligned} a &= a_0\ a_1\ a_2 \ldots a_{q-1} a_q\ a_{q+1} \ldots a_{r-1} \\ a_{(q)} &= a_{r-q} a_{r-q+1} a_{r-q+2} \ldots a_{r-1}\ a_0\ a_1 \ldots a_{r-q-1} \\ a_{(-q)} &= a_q\ a_{q+1}\ a_{q+2} \ldots a_{2q-1} a_{2q} a_{2q+1} \ldots a_{q-1} \end{aligned}$$

Klar ist $\kappa(a,a_{(q)}) = \kappa(a,a_{(-q)})$.

Definition. Für einen Text $a \in \Sigma^*$ und eine natürliche Zahl $q \in \mathbf{N}$ heißt

$$\kappa_q(a) := \kappa(a,a_{(q)})$$

der q-te **Autokoinzidenzindex** von a.

Achtung: In der Literatur ist diese Bezeichnung nicht üblich. Eine übliche Bezeichnung gibt es für diese Größe nicht.

Beispiele.

Eigenschaften

Der q-te Autokoinzidenzindex definiert eine Abbildung

$$\kappa_q : \Sigma^* \to \mathbf{Q}.$$

Die konstante Abbildung κ_0 ist natürlich uninteressant.

Klar ist $\kappa_q(a) = \kappa_{r-q}(a)$ für $a \in \Sigma^r$ und $0 < q < r$.

Anwendung

Betrachten wir nun einen periodisch polyalphabetisch verschlüsselten Text. Bei der Bestimmung von $\kappa_q(a)$ wird im allgemeinen die Verschiebung q kein Vielfaches der Periode l sein. Bei der Zählung der Koinzidenzen treffen also Buchstaben zusammen, die im wesentlichen unabhängig

voneinander monoalphabetisch verschlüsselt sind. Wir werden also nach I.3.1 ein Ergebnis $\kappa_q(f(a)) \approx 1/n$ erwarten.

Ist allerdings $l|q$, so haben wir die Situation

$$\begin{array}{cccccccc} \sigma_0 a_0 & \sigma_1 a_1 & \ldots & \sigma_0 a_q & \sigma_1 a_{q+1} & \ldots \\ & & & \sigma_0 a_0 & \sigma_1 a_1 & \ldots \end{array}$$

wo also die untereinander stehenden Zeichen mit der gleichen monoalphabetischen Substitution verschlüsselt sind und somit genau dann übereinstimmen, wenn sie im Klartext übereinstimmten. Wir erwarten jetzt also, dass $\kappa_q(f(a))$ in der Nähe der für die Sprache typischen Zeichenkoinzidenz liegt.

Genauer ist also für eine polyalphabetische Verschlüsselung f der Periode l folgendes zu erwarten:

1. Ist l kein Teiler von q und r-q, so ist $\kappa_q(f(a)) \approx 1/n$.
2. Ist aber l Teiler von q und q klein gegenüber r, so ist $\kappa_q(f(a)) \approx \kappa_q(a)$ ungefähr die typische Zeichenkoinzidenz.

Das ist die **zweite Anwendung der Koinzidenzbestimmung**: Erkennung der Periode einer polyalphabetischen Substitution an den Autokoinzidenzindizes. Im Vergleich zur Parallelstellenanalyse berücksichtigt ein Autokoinzidenzindex auch Wiederholungen der Länge 1 im Abstand q; d. h., die Information, die die Periode im Geheimtext hinterlässt, wird wesentlich besser ausgenützt als bei der Parallestellensuche nach KASISKI. (»Statt nur einem oberflächlichen Foto wird ein Röntgenbild angefertigt.«) Man kann hoffen, dass sich die Periode einer polyalphabetischen Chiffre deutlich verrät.

Um eine Vorstellung von der Bedeutung dieser noch vagen Aussagen zu erhalten, rechnen wir als nächstes ein Beispiel durch; das dazu benützte Perl-Programm steht hier [online-Aufruf].

Autor: Klaus Pommerening, 23. Februar 2000; letzte Änderung: 23. November 2004.
E-Mail an `Pommerening »AT« imbei.uni-mainz.de`.