Odšifriranje Vigenèrejeve šifre

Test Friedericha Kasiskega (1863):
(in Charles Babbage-a 1854)

poiščemo dele tajnopisa $\boldsymbol{y} = y_1 y_2 \dots y_n$, ki so identični in zabeležimo razdalje $d_1, d_2, \dots$ med njihovimi začetki. Predpostavimo, da iskani m deli največji skupni delitelj teh števil.

Naj bo $d = n/m$. Elemente tajnopisa $\boldsymbol{y}$ zapišemo po stolpcih v $(m \times d)$-razsežno matriko. Vrstice označimo z $\boldsymbol{y_i}$, tj.

$$\boldsymbol{y_i} = y_i \, y_{m+i} \, y_{2m+i} \, \dots$$

Indeks naključja (William Friedman, 1920):

Za zaporedje $\boldsymbol{x} = x_1 x_2 \dots x_d$ je **indeks naključja** (angl. index of coincidence, oznaka $I_c(\boldsymbol{x})$) **verjetnost**, da sta naključno izbrana elementa zaporedja $\boldsymbol{x}$ enaka.

Če so $f_0, f_1, \dots, f_{25}$ frekvence črk $A, B, \dots, Z$ v zaporedju $\boldsymbol{x}$, je

$$I_c(\boldsymbol{x}) = \frac{\sum_{i=0}^{25} \binom{f_i}{2}}{\binom{d}{2}} = \sum_{i=0}^{25} \frac{f_i(f_i-1)}{d(d-1)}.$$

Če so p_i pričakovane verjetnosti angleških črk, potem je

$$I_c(\boldsymbol{x}) \approx \sum_{i=0}^{25} p_i^2 = 0.065.$$

Za povsem naključno zaporedje velja

$$I_c(\boldsymbol{x}) \approx 26 \left(\frac{1}{26}\right)^2 = \frac{1}{26} = 0.038.$$

Ker sta števili .065 in .038 dovolj narazen, lahko s to metodo najdemo dolžino ključa

(ali pa potrdimo dolžino, ki smo jo uganili s testom Kasiskega).

Za podzaporedje $\boldsymbol{y_i}$ in $0 \leq g \leq 25$ naj bo

$$M_g(\boldsymbol{y_i}) = \sum_{i=0}^{25} p_i \frac{f_{i+g}}{d}.$$

Če je $g = k_i$, potem pričakujemo

$$M_g(\boldsymbol{y_i}) \approx \sum_{i=0}^{25} p_i^2 = 0.065$$

Za $g \neq k_i$ je običajno M_g bistveno manjši od 0.065.

Torej za vsak $1 \leq i \leq m$ in $0 \leq g \leq 25$ tabeliramo vrednosti M_g, nato pa v tabeli za vsak $1 \leq i \leq m$ poiščemo tiste vrednosti, ki so blizu 0.065.

Ustrezni g-ji nam dajo iskane zamike $k_1, k_2, \dots, k_m$.

Odšifriranje Hillove šifre

Predpostavimo, da je nasprotnik določil m, ki ga uporabljamo, ter se dokopal do m različnih parov m-teric (2. stopnja – poznan čistopis):

$$x_j = (x_{1,j}, x_{2,j}, \dots, x_{m,j}), \;\; y_j = (y_{1,j}, y_{2,j}, \dots, y_{m,j}),$$

tako da je $y_j = e_K(x_j)$ za $1 \leq j \leq m$.

Za matriki $X = (x_{i,j})$ in $Y = (y_{i,j})$ dobimo matrično enačbo $Y = XK$.

Če je matrika X obrnljiva, je $K = YX^{-1}$.

Za Hillovo šifro lahko uporabimo tudi 1. stopnjo napada (samo tajnopis), glej nalogo 1.25.

Koliko ključev imamo na voljo v primeru Hillove šifre? Glej nalogo 1.12.

Za afino-Hillovo šifro glej nalogo 1.24.

Tokovne šifre

Naj bo $x_1 x_2 \dots$ čistopis.

Doslej smo obravnavali kriptosisteme z enim samim ključem in tajnopis je imel naslednjo obliko.

$$\mathbf{y} = y_1 y_2 \dots = e_K(x_1) e_K(x_2) \dots$$

Taki šifri pravimo **bločna šifra** (angl. block cipher).

Posplošitev: iz enega ključa $K \in \mathcal{K}$ napravimo zaporedje (tok) ključev. Naj bo f_i funkcija, ki generira i-ti ključ:

$$z_i = f_i(K, x_1, \dots, x_{i-1}).$$

Z njim izračunamo:

$$y_i = e_{z_i}(x_i) \qquad \text{in} \qquad x_i = d_{z_i}(y_i).$$

Bločna šifra je poseben primer tokovne šifre (kjer je $z_i = K$ za vse $i \geq 1$).

Sinhrona tokovna šifra *je sedmerica* $(\mathcal{P},\mathcal{C},\mathcal{K},\mathcal{L},\mathcal{F},\mathcal{E},\mathcal{D})$ *za katero velja:*

1. $\mathcal{P}$ *je končna množica možnih* **čistopisov**,
2. $\mathcal{C}$ *je končna množica možnih* **tajnopisov**,
3. $\mathcal{K}$ *je končna množica možnih* **ključev**,
4. $\mathcal{L}$ *je končna množica tokovne* **abecede**,
5. $\mathcal{F} = (f_1, f_2, \dots)$ *je generator toka ključev:*
$$f_i : \mathcal{K} \times \mathcal{P}^{i-1} \longrightarrow \mathcal{L} \quad \text{za } i \geq 1$$
6. *Za vsak ključ* $z \in \mathcal{L}$ *imamo šifrirni* $(e_z \in \mathcal{E})$ *in odšifrirni* $(d_z \in \mathcal{D})$ *postopek, tako da je* $d_z(e_z(x)) = x$ *za vsak* $x \in \mathcal{P}$.

Za šifriranje čistopisa $x_1x_2\dots$ zaporedno računamo
$$z_1, y_1, z_2, y_2, \dots,$$
za odšifriranje tajnopisa $y_1y_2\dots$ pa zaporedno računamo
$$z_1, x_1, z_2, x_2, \dots.$$

Tokovna šifra je **periodična** s periodo d kadar, je $z_{i+d} = z_i$ za vsak $i \geq 1$

(poseben primer: Vigenèrejeva šifra).

Začnimo s ključi $(k_1, \dots, k_m)$ in naj bo $z_i = k_i$ za $i = 1, \dots, m$.
Definiramo linearno rekurzijo stopnje m:
$$z_{i+m} = z_i + \sum_{j=1}^{m-1} c_j\, z_{i+j} \mod 2,$$
kjer so $c_1, \dots, c_{m-1} \in \mathbb{Z}_2$ vnaprej določene konstante.

Za ustrezno izbiro konstant $c_1, \dots, c_{m-1} \in \mathbb{Z}_2$ in neničelen vektor $(k_1, \dots, k_m)$ lahko dobimo tokovno šifro s periodo $2^m - 1$.

Hitro lahko generiramo tok ključev z uporabo **LFSR** (**Linear Feedback Shift Register**).
V pomičnem registru začnemo z vektorjem
$$(k_1, \dots, k_m).$$
Nato na vsakem koraku naredimo naslednje:

1. k_1 dodamo toku ključev (za XOR),
2. $k_2, \dots, k_m$ pomaknemo za eno v levo,
3. 'nov' ključ k_m izračunamo z
$$\sum_{j=0}^{m-1} c_j\, k_{j+1} \quad \text{(to je "linear feedback")}.$$

Primer:

$c_0 = 1, c_1 = 1, c_2 = 0, c_3 = 0,$

torej je $k_{i+4} = k_i + k_{i+1}$.

Izberimo $k_0 = 1, k_1 = 0, k_2 = 1, k_3 = 0$.

Potem je $k_4 = 1, k_5 = 1, k_6 = 0, \dots$.

Naj bo $\boldsymbol{k} = (k_0, k_1, k_2, k_3)^t$ in
$$A := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$
Torej je $A(\boldsymbol{k}) = (k_1, k_2, k_3, k_4)^t$,
$A^2(\boldsymbol{k}) = A(k_1, k_2, k_3, k_4)^t = (k_2, k_3, k_4, k_5)^t$
...
$A^i(\boldsymbol{k}) = (k_i, k_{i+1}, k_{i+2}, k_{i+3})^t$.

Najdaljša možna perioda je 15.

Enkrat dobimo:
$$A^i(\boldsymbol{k}) = A^j(\boldsymbol{k})$$
in ker je A obrnljiva
$$A^{i-j}(\boldsymbol{k}) = \boldsymbol{k}$$
Karakteristični polinom matrike A je
$$f(x) = 1 + x + x^4.$$
Ker je $f(x)$ nerazcepen, je $f(x)$ tudi minimalni polinom matrike A.

Red matrike A je najmanjše naravno število s, tako da je $A^s = I$. Naj bo e najmanjše naravno število, tako da $f(x) \mid (x^e - 1)$. Potem je $e = s$.
$$1 + x^{15} = (x+1)(x^2+x+1)(x^4+x+1)(x^4+x^3+1)(x^4+x^3+x^2+x+1).$$
Splošno: če hočemo, da nam rekurzija stopnje m da periodo $2^m - 1$, potem si izberemo nerazcepen f.

Analiza je neodvisna od začetnega neničelnega vektorja.

Kriptoanaliza LFSR tokovne šifre:
uporabimo lahko poznan čistopis, glej nalogo 1.27.

2. poglavje

Shannonova teorija

- Popolna varnost
- Entropija
- Lastnosti entropije
- Ponarejeni ključi
  in enotska razdalja
- Produktne šifre

Popolna varnost

Omenimo nekaj osnovnih principov za študij varnosti nekega kriptosistema:

- **računska varnost**,
- **brezpogojna varnost**,
- **dokazljiva varnost**.

Kriptosistem je **računsko varen**, če tudi najboljši algoritem za njegovo razbitje potrebuje vsaj N operacij, kjer je N neko konkretno in zelo veliko število.

Napadalec (Oskar) ima na razpolago 18 Crayev, 4000 Pentium PC-jev in 200 DEC Alpha mašin (Oskar je "računsko omejen").

Kriptosistem je **dokazljivo varen** (angl. provable secure), če lahko pokažemo, da se njegova varnost zreducira na varnost kriptosistema, ki je zasnovan na dobro preštudiranem problemu.

Ne gre torej za absolutno varnost temveč *relativno varnost*.

Gre za podobno strategijo kot pri dokazovanju, da je določen problem *NP-poln* (v tem primeru dokažemo, da je dani problem vsaj tako težak kot nekdrugi znani NP-poln problem, ne pokažemo pa, da je absolutno računsko zahteven).

Kriptosistem je **brezpogojno varen**, kadar ga napadalec ne more razbiti, tudi če ima na voljo neomejeno računsko moč.

Seveda je potrebno povedati tudi, kakšne vrste napad imamo v mislih. Spomnimo se, da zamične, substitucijske in Vigenère šifre niso varne pred napadom s poznanim tajnopisom (če imamo na voljo dovolj tajnopisa).

Razvili bomo teorijo kriptosistemov, ki so brezpogojno varni pri napadu s poznanim tajnopisom. Izkaže se, da so vse tri šifre brezpogojno varne, kadar zašifriramo le en sam element čistopisa.

Glede na to, da imamo pri brezpogojni varnosti na voljo neomejeno računsko moč, je ne moremo študirati s pomočjo teorije kompleksnosti, temveč s teorijo verjetnosti.

Naj bosta X in Y slučajni spremenljivki,
naj bo $p(x) := P(X = x)$, $p(y) := P(Y = y)$ in
$p(x \cap y) := P((X=x) \cap (Y=y))$ produkt dogodkov.

Slučajni spremenljivki X in Y sta **neodvisni**, če in samo, če je $p(x \cap y) = p(x)p(y)$ za vsak $x \in X$ in $y \in Y$.

Omenimo še zvezo med pogojno verjetnostjo in pa verjetnostjo produkta dveh dogodkov oziroma **Bayesov izrek o pogojni verjetnosti**:

$$p(x \cap y) = p(x/y)p(y) = p(y/x)p(x),$$

iz katerega sledi, da sta slučajni spremenljivki X in Y neodvisni, če in samo, če je $p(x/y) = p(x)$ za vsak x in y.

Privzemimo, da vsak ključ uporabimo za največ eno šifriranje, da si Anita in Bojan izbereta ključ K z neko fiksno verjetnostno porazdelitvijo $p_{\mathcal{K}}(K)$ (pogosto enakomerno porazdelitvijo, ni pa ta nujna) in naj bo $p_{\mathcal{P}}(x)$ verjetnost čistopisa x.

Končno, predpostavimo, da sta izbira čistopisa in ključa neodvisna dogodka.

Porazdelitvi $\mathcal{P}$ in $\mathcal{K}$ inducirata verjetnostno porazdelitev na $\mathcal{C}$. Za množico vseh tajnopisov za ključ K

$$C(K) = \{e_K(x) \,|\, x \in \mathcal{P}\}$$

velja

$$p_{\mathcal{C}}(y) = \sum_{\{K \,|\, y \in C(K)\}} p_{\mathcal{K}}(K)\, p_{\mathcal{P}}(d_k(y))$$

in

$$P(Y = y/X = x) = \sum_{\{K \,|\, x = d_K(y)\}} p_{\mathcal{K}}(K).$$

Sedaj lahko izračunamo pogojno verjetnost $p_{\mathcal{P}}(x/y)$, tj. verjetnost, da je x čistopis, če je y tajnopis

$$P(X = x/Y = y) = \frac{p_{\mathcal{P}}(x) \times \displaystyle\sum_{\{K \,|\, x = d_K(y)\}} p_{\mathcal{K}}(K)}{\displaystyle\sum_{\{K \,|\, y \in C(K)\}} p_{\mathcal{K}}(K)\, p_{\mathcal{P}}(d_k(y))}$$

in opozorimo, da jo lahko izračuna vsakdo, ki pozna verjetnostni porazdelitvi $\mathcal{P}$ in $\mathcal{K}$.

Primer: $\mathcal{P} = \{a, b\}$ in $\mathcal{K} = \{K_1, K_2, K_3\}$:

$$p_{\mathcal{P}}(a) = 1/4 \text{ in } p_{\mathcal{P}}(b) = 3/4.$$

$$p_{\mathcal{K}}(K_1) = 1/2 \text{ in } p_{\mathcal{K}}(K_2) = p_{\mathcal{K}}(K_3) = 1/4.$$

Enkripcija pa je definirana z $e_{K_1}(a) = 1$, $e_{K_1}(b) = 2$; $e_{K_2}(a)=2$, $e_{K_2}(b)=3$; $e_{K_3}(a)=3$, $e_{K_3}(b)=4$.

Potem velja

$$p_{\mathcal{C}}(1)=\frac{1}{8},\;\; p_{\mathcal{C}}(2)=\frac{7}{16},\;\; p_{\mathcal{C}}(3)=\frac{1}{4},\;\; p_{\mathcal{C}}(3)=\frac{3}{16}.$$

$$p_{\mathcal{P}}(a/1)=1,\; p_{\mathcal{P}}(a/2)=\frac{1}{7},\; p_{\mathcal{P}}(a/3)=\frac{1}{4},\; p_{\mathcal{P}}(a/4)=0.$$

Šifra $(\mathcal{P}, \mathcal{K}, \mathcal{C})$ je **popolnoma varna**, če je

$$P(X = x/Y = y) = p_{\mathcal{P}}(x) \quad \text{za vse } x \in \mathcal{P} \text{ in } y \in \mathcal{C},$$

tj. "končna" verjetnost, da smo začeli s tajnopisom x pri danem čistopisu y, je identična z "začetno" verjetnostjo čistopisa x.

V prejšnjem primeru je ta pogoj zadoščen samo v primeru $y = 3$, ne pa tudi v preostalih treh.

Izrek 1. *Če ima vseh 26 ključev pri zamični šifri enako verjetnost $1/26$, potem je za vsako verjetnostno porazdelitev čistopisa zamična šifra popolnoma varna.*

Dokaz: $\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbb{Z}_{26}$, $e_K(x) = x + K \bmod 26$:

$$p_{\mathcal{C}}(y) = \frac{1}{26} \sum_{K \in \mathbb{Z}_{26}} p_{\mathcal{P}}(y - K) = \frac{1}{26},$$

$$P(Y = y/X = x) = p_{\mathcal{K}}(y - x \bmod 26)) = \frac{1}{26}. \qquad \blacksquare$$

Torej lahko zaključimo, da zamične šifre ne moremo razbiti, če za vsak znak čistopisa uporabimo nov, naključno izbran ključ.

Sedaj pa preučimo popolno varnost na splošno. Pogoj

$$P(X = x/Y = y) = p_{\mathcal{P}}(x) \quad \text{za vse } x \in \mathcal{P} \text{ in } y \in \mathcal{C}$$

je ekvivalenten pogoju

$$P(Y = y/X = x) = p_{\mathcal{C}}(y) \quad \text{za vse } x \in \mathcal{P} \text{ in } y \in \mathcal{C}.$$

Privzemimo (BŠS), da je $p_{\mathcal{C}}(y) > 0$ za vse $y \in \mathcal{C}$. Ker je $P(Y = y/X = x) = p_{\mathcal{C}}(y) > 0$ za fiksen $x \in \mathcal{P}$ in za vsak $y \in \mathcal{C}$, za vsak tajnopis $y \in \mathcal{C}$ obstaja vsaj en ključ K, da je $e_K(x) = y$ in zato velja $|\mathcal{K}| \geq |\mathcal{C}|$.

Za vsako simetrično šifro velja $|\mathcal{C}| \geq |\mathcal{P}|$, saj smo privzeli, da je šifriranje injektivno.

V primeru enakosti (v obeh neenakostih) je Shannon karakteriziral popolno varnost na naslednji način:

Izrek 2. *Naj bo $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ simetrična šifra za katero velja $|\mathcal{K}| = |\mathcal{C}| = |\mathcal{P}|$. Potem je le-ta popolnoma varna, če in samo, če je vsak ključ uporabljen z enako verjetnostjo $1/|\mathcal{K}|$ ter za vsak čistopis x in za vsak tajnopis y obstaja tak ključ K, da je $e_K(x) = y$.*

Dokaz: ($\Longrightarrow$) Ker je $|\mathcal{K}| = |\mathcal{C}|$, sledi, da za vsak čistopis $x \in \mathcal{P}$ in za vsak tajnopis $y \in \mathcal{C}$ obstaja tak ključ K, da je $e_K(x) = y$.

Naj bo $n = |\mathcal{K}|$, $\mathcal{P} = \{x_i \mid 1 \leq i \leq n\}$ in naj za fiksen tajnopis y označimo ključe iz $\mathcal{K}$ tako, da je $e_{K_i}(x_i) = y$ za $i \in [1..n]$. Po Bayesovem izreku velja

$$P(X = x_i/Y = y) = \frac{P(Y = y/x = x_i)\, p_{\mathcal{P}}(x_i)}{p_{\mathcal{C}}(y)} = \frac{p_{\mathcal{K}}(K_i)\, p_{\mathcal{P}}(x_i)}{p_{\mathcal{C}}(y)}.$$

Če je šifra popolnoma varna, velja $P(X = x_i/Y = y) = p_{\mathcal{P}}(x_i)$, torej tudi $p_{\mathcal{K}}(K_i) = p_{\mathcal{C}}(y)$, kar pomeni, da je vsak ključ uporabljen z enako verjetnostjo $p_{\mathcal{C}}(y)$ in zato $p_{\mathcal{K}}(K) = 1/|\mathcal{K}|$.

Dokaz obrata poteka na podoben način kot v prejšnjem izreku. $\blacksquare$

Najbolj znana realizacija popolne varnosti je **Vernamov enkratni ščit**, ki ga je leta 1917 patentiral Gilbert Vernam za avtomatizirano šifriranje in odšifriranje telegrafskih sporočil.

Naj bo $\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_2)^n$, $n \in \mathbb{N}$,

$$e_K(x) = x \text{ XOR } K,$$

odšifriranje pa je identično šifriranju.

Shannon je prvi po 30-ih letih dokazal, da ta sistem res ne moremo razbiti.

Slabi strani te šifre sta $|\mathcal{K}| \geq |\mathcal{P}|$ in dejstvo, da moramo po vsaki uporabi zamenjati ključ.

Entropija

Doslej nas je zanimala popolna varnost in smo se omejili na primer, kjer uporabimo nov ključ za vsako šifriranje.

Sedaj pa nas zanimata šifriranje vse več in več čistopisa z istim ključem ter verjetnost uspešnega napada z danim tajnopisom in neomejenim časom.

Leta 1948 je Shannon vpeljal v teorijo informacij *entropijo*, tj. matematično mero za informacije oziroma negotovosti in jo izrazil kot funkcijo verjetnostne porazdelitve.

Naj bo X slučajna spremenljivka s končno zalogo vrednosti in porazdelitvijo $p(X)$.

Kakšno informacijo smo pridobili, ko se je zgodil dogodek glede na porazdelitev $p(X)$

oziroma ekvivalentno,

če se dogodek še ni zgodil, kolikšna je negotovost izida?

To količino bomo imenovali **entropija** spremenljivke X in jo označili s $H(X)$.

Primer: metanje kovanca, $p(\text{cifra}) = p(\text{grb}) = 1/2$.

Smiselno je reči, da je entropija enega meta en bit.

Podobno je entropija n-tih metov n, saj lahko rezultat zapišemo z n biti.

Še en primer: slučajna spremenljivka X

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Najbolj učinkovito zakodiranje izidov je x_1 z 0, x_2 z 10 in x_3 z 11, povprečje pa je

$$\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 2 \ = \ 3/2.$$

Vsak dogodek, ki se zgodi z verjetnostjo 2^{-n}, lahko zakodiramo z n biti.

Posplošitev: dogodek, ki se zgodi z verjetnostjo p, lahko zakodiramo s približno $-\log_2 p$ biti.

Naj bo X slučajna spremenljivka s končno zalogo vrednosti in porazdelitvijo

$$p(X) = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}.$$

Potem **entropijo porazdelitve** $p(X)$ definiramo s

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i = -\sum_{i=1}^{n} p(X{=}x_i) \log_2 p(X{=}x_i).$$

Za $p_i = 0$ količina $\log_2 p_i$ ni definirana, zato seštevamo samo po neničelnih p_i (tudi $\lim_{x \to 0} x \log_2 x = 0$).

Lahko bi izbrali drugo logaritemsko bazo, a bi se entropija spremenila le za konstantni faktor.

Če je $p_i = 1/n$ za $1 \leq i \leq n$, potem je $H(X) = \log_2 n$.

Velja $H(X) \geq 0$, enačaj pa velja, če in samo, če je $p_i = 1$ za nek i in $p_j = 0$ za $j \neq i$.

Sedaj pa bomo študirali entropijo različnih komponent simetrične šifre: $H(K)$, $H(P)$, $H(C)$.

Za primer $\mathcal{P} = \{a, b\}$ in $\mathcal{K} = \{K_1, K_2, K_3\}$:

$$p_{\mathcal{P}}(a) = 1/4 \ \text{ in } \ p_{\mathcal{P}}(b) = 3/4.$$

$$p_{\mathcal{K}}(K_1) = 1/2 \ \text{ in } \ p_{\mathcal{K}}(K_2) = p_{\mathcal{K}}(K_3) = 1/4$$

izračunamo

$$H(P) = -\frac{1}{4} \log_2 \frac{1}{4} - \frac{3}{4} \log_2 \frac{3}{4} = 2 - \frac{3}{4} \log_2 3 \approx .81 \ .$$

in podobno $H(K) = 1.5$ ter $H(C) \approx 1.85$.

Lastnosti entropije

Realna funkcija f je **(striktno) konkavna** na intervalu I, če za vse (različne) $x, y \in I$ velja

$$f\left(\frac{x+y}{2}\right) \ (>) \ \geq \ \frac{f(x) + f(y)}{2}.$$

Jensenova neenakost: če je f zvezna in striktno konkavna funkcija na intervalu I in $\sum_{i=1}^{n} a_i = 1$ za $a_i > 0$, $1 \leq i \leq n$, potem je

$$f\left(\sum_{i=1}^{n} a_i x_i\right) \geq \sum_{i=1}^{n} a_i f(x_i),$$

enakost pa velja, če in samo, če je $x_1 = x_2 = \cdots = x_n$.

Izrek 3. *$H(X) \leq \log_2 n$, enakost pa velja, če in samo, če je $p_1 = p_2 = \cdots = p_n = 1/n$.*

Izrek 4. *$H(X, Y) \leq H(X) + H(Y)$, enakost pa velja, če in samo, če sta X in Y neodvisni spremenljivki.*

Dokaz izreka 4: Naj bo

$$p(X) = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix}, \ p(Y) = \begin{pmatrix} y_1 & y_2 & \dots & y_n \\ q_1 & q_2 & \dots & q_n \end{pmatrix}$$

in $r_{ij} = p((X{=}x_i) \cap (Y{=}y_j))$ za $i \in [1..m]$, $j \in [1..n]$.

Potem za $i \in [1..m]$ in $j \in [1..n]$ velja

$$p_i = \sum_{j=1}^{n} r_{ij} \quad \text{in} \quad q_j = \sum_{i=1}^{m} r_{ij}$$

ter

$$H(X) + H(Y) = -\sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} \log_2 p_i q_j$$

in

$$H(X, Y) - H(X) - H(Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} \log_2 \frac{p_i q_j}{r_{ij}}$$

$$(\text{Jensen}) \leq \log_2 \sum_{i=1}^{m} \sum_{j=1}^{n} p_i q_j = \log_2 1 = 0.$$

Enakost velja, če in samo, če je $p_i q_j / r_{ij} = c$ za $i \in [1..m]$ in $j \in [1..n]$.

Upoštevajmo še

$$\sum_{j=1}^{n}\sum_{i=1}^{m} r_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m} p_i q_j = 1$$

in dobimo $c = 1$ oziroma za vse i in j

$$p((X = x_i) \cap (Y = y_j)) = p(X = x_i)\, p(Y = y_j),$$

kar pomeni, da sta spremenljivki X in Y neodvisni. ■

Za slučajni spremenljivki X in Y definiramo **pogojni entropiji**

$$H(X/y) = -\sum_{x} p(x/y) \log_2 p(x/y)$$

in

$$H(X/Y) = -\sum_{y}\sum_{x} p(y) p(x/y) \log_2 p(x/y).$$

Le-ti merita povprečno informacijo spremenljivke X, ki jo odkrijeta y oziroma Y.

Izrek 5. $H(X,Y) = H(Y) + H(X/Y)$.

Dokaz: Po definiciji je $P(X = x_i / Y = y_j) = r_{ij}/q_j$ in $H(Y) + H(X/Y) =$

$$= -\sum_{j=1}^{n}\sum_{i=1}^{m} r_{ij} \log_2 q_j - \sum_{j=1}^{n}\sum_{i=1}^{m} q_j\, r_{ij}/q_j\, \log_2 r_{ij}/q_j \quad ■$$

Iz izrekov 4 in 5 sledi:

Posledica 6. *$H(X/Y) \le H(X)$, enakost pa velja, če in samo, če sta X in Y neodvisni spremenljivki.*

Ponarejeni ključi in enotska razdalja

Pogojna verjetnost $H(K/C)$ meri, koliko informacije o ključu je odkrito s tajnopisom.

Izrek 7. *Naj bo $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ simetrična šifra. Potem velja $H(K/C) = H(K) + H(P) - H(C)$.*

Dokaz: Velja $H(K,P,C) = H(C/(K,P)) + H(K,P)$. Ker ključ in čistopis natanko določata tajnopis, je $H(C/K,P) = 0$.

Ker sta P in K neodvisni spremenljivki, dobimo $H(K,P,C) = H(P) + H(K)$ in podobno tudi $H(K,P,C) = H(K,C)$ ter uporabimo še izrek 5. ■

Napadalec privzame, da je čistopis "naravni" jezik (npr. angleščina) in na ta način odpiše mnoge ključe. Vseeno pa lahko ostane še mnogo ključev (med katerimi je le en pravi), ki jih bomo, razen pravega ključa, imenovali **ponarejeni** (angl. spurious).

Naš cilj bo oceniti število ponarejenih ključev.

Naj bo H_L mera povprečne informacije na črko (angl. per letter) v "smiselnem" čistopisu (sledi bolj natančna definicija).

Če so vse črke enako verjetne, je

$$H_L = \log_2 26 \approx 4.70.$$

Kot aproksimacijo *prvega reda* bi lahko vzeli $H(P)$. V primeru angleškega jezika dobimo $H(P) \approx 4.19$.

Tudi zaporedne črke v jeziku niso neodvisne, njihove korelacije pa zmanjšajo entropijo. Za aproksimacijo *drugega reda* bi lahko izračunali entropijo porazdelitve parov črk in potem delili z dve, kajti H_L meri entropijo jezika L na črko.

V splošnem, naj bo P^n slučajna spremenljivka, katere verjetnostna porazdelitev je enaka verjetnostni porazdelitvi n-teric v čistopisu.

Potem je **entropija za naravni jezik L** definirana s

$$H_L = \lim_{n \to \infty} \frac{H(P^n)}{n},$$

odvečnost jezika L pa z

$$R_L = 1 - \frac{H_L}{\log_2 |\mathcal{P}|}.$$

H_L meri entropijo jezika L na črko.

Entropija naključnega jezika je $\log_2 |\mathcal{P}|$.

$R_L \in [0, 1)$ meri kvocient "odvečnih znakov" in je 0 v primeru naključnega jezika.

Za angleški jezik je $H(P^2)/2 \approx 3.90$.

Empirični rezultati kažejo, da je $1.0 \le H_L \le 1.5$.

Če ocenimo H_L z 1.25, potem je $R_L \approx .75$, kar pomeni, da je angleščina 75% odvečna

(tj. tekst bi lahko zakodirali le z 1/4 prvotnega teksta).

Podobno kot P^n definiramo še C^n in za $y \in C^n$ še

$$K(y) = \{K \in \mathcal{K} \,|\, \exists x \in P^n,\, p_{\mathcal{P}^n}(x) > 0,\, e_K(x) = y\},$$

tj. $K(y)$ je množica ključev, za katere je y smiselno šifriranje čistopisa dolžine n,

tj. množica verjetnih ključev, za katere je y tajnopis.

Matematično upanje ponarejenih ključev je torej

$$\overline{s}_n = \sum_{y \in \mathcal{C}^n} p(y)(|K(y)| - 1) = \sum_{y \in \mathcal{C}^n} p(y)|K(y)| \ - \ 1.$$

Izrek 8. *Če je $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ šifra za katero je $|\mathcal{C}| = |\mathcal{P}|$ in so vsi ključi med seboj enakovredni, potem za tajnopis z n znaki (n je dovolj velik) in za matematično upanje ponarjenih ključev $\overline{s}_n$ velja*

$$\overline{s}_n \geq \frac{|\mathcal{K}|}{|\mathcal{P}|^{nR_L}} - 1.$$

Dokaz: Iz izreka 7 sledi

$$H(K/C^n) = H(K) + H(P^n) - H(C^n).$$

Poleg ocene $H(C^n) \leq n \log_2 |\mathcal{C}|$ velja za dovolj velike n tudi ocena $H(P^n) \approx nH_L = n(1 - R_L) \log_2 |\mathcal{P}|$.

Za $|\mathcal{C}| = |\mathcal{P}|$ dobimo

$$H(K/C^n) \geq H(K) - nR_L \log_2 |\mathcal{P}|.$$

Ocenjeno entropijo povežemo še s ponarejenimi ključi

$$H(K/C^n) = \sum_{y \in \mathcal{C}^n} p(y) H(K/y) \leq \sum_{y \in \mathcal{C}^n} p(y) \log_2 |K(y)|$$

$$\leq \log_2 \sum_{y \in \mathcal{C}^n} p(y)|K(y)| = \log_2(\overline{s}_n + 1). \ \blacksquare$$

Desna stran neenakosti v zadnjem izreku gre z večanjem števila n eksponentno proti 0 (to ni limita, števila $|\mathcal{K}|$, $|\mathcal{P}|$ in R_L so fiksna, število $|\mathcal{K}|$ pa je običajno veliko v primerjavi s $|\mathcal{P}|^{R_L} > 1$).

Enotska razdalja simetrične šifre je tako število n, označeno z n_0, za katerega postane matematično upanje ponarejenih ključev nič, tj. povprečna dolžina tajnopisa, ki jo napadalec potrebuje za računanje ključa pri neomejenem času.

Velja $$n_0 \approx \frac{\log_2 |\mathcal{K}|}{R_L \log_2 |\mathcal{P}|}.$$

V primeru zamenjalnega tajnopisa sta $|\mathcal{P}| = 26$ in $|\mathcal{K}| = 26!$. Če vzamemo $R_L = .75$, potem je enotska razdalja

$$n_0 \approx \frac{88.4}{.75 \times 4.7} \approx 25 \ .$$

Produktne šifre

Še ena Shannonova ideja v članku iz leta 1949 igra danes pomembno vlogo, predvsem pri simetričnih šifrah.

Zanimali nas bodo šifre, za katere $\mathcal{C} = \mathcal{P}$, tj. **endomorfne šifre**.

Naj bosta $S_i = (\mathcal{P}, \mathcal{P}, \mathcal{K}_i, \mathcal{E}_i, \mathcal{D}_i)$, $i = 1, 2$, endomorfni simetrični šifri. Potem je **produkt** sistemov S_1 in S_2, označen s $S_1 \times S_2$, definiran s

$$(\mathcal{P}, \mathcal{P}, \mathcal{K}_1 \times \mathcal{K}_2, \mathcal{E}, \mathcal{D})$$

ter

$$e_{(K_1,K_2)}(x) = e_{K_2}(e_{K_1}(x))$$

in

$$d_{(K_1,K_2)}(y) = d_{K_1}(e_{K_2}(y)).$$

Njegova verjetnostna porazdelitev pa naj bo

$$p_{\mathcal{K}}(K_1, K_2) = p_{\mathcal{K}_1}(K_1) \times p_{\mathcal{K}_2}(K_2),$$

tj. ključa K_1 in K_2 izberemo neodvisno.

Če sta M in S zaporedoma multiplikativni tajnopis in zamični tajnopis, potem je $M \times S$ afin tajnopis. Malce težje je pokazati, da je tudi tajnopis $S \times M$ afin tajnopis. Ta dva tajnopisa torej **komutirata**.

Vsi tajnopisi ne komutirajo, zato pa je produkt asociativna operacija:

$$(S_1 \times S_2) \times S_3 = S_1 \times (S_2 \times S_3).$$

Če je $(S \times S \ =) \ S^2 = S$, pravimo, da je sistem **idempotenten**.

Zamični, zamenjalni, afin, Hillov, Vigenèrov in permutacijski tajnopisi so vsi idempotentni.

Če simetrična šifra ni idempotentna, potem se zna zgoditi, da z njeno iteracijo za večkrat povečamo varnost. Na tem so zasnovani **DES** in mnoge druge simetrične šifre.

Če sta simetrični šifri S_1 in S_2 idempotentni in obenem še komutirata, potem se ni težko prepričati, da je tudi produkt $S_1 \times S_2$ idempotentna simetrična šifra.