Kryptologia

T. Katriňák

October 4, 2010

Historia (1. lekcia)

Kryptologia je veda o utajovani dat a tiez o tom, ako mozno utajene data "odtajnit". Preto kryptologia sa sklada z dvoch velkych casti: z **kryptografie**, t.j. vede o utajovani dat pomocou sifrovania (kodovania) a z **kryptoanalyzy**, t.j. z umenia (?) ci vedy, ktore nam pomozu zasifrovane data desifrovat.

Historicke priklady: 1. Grecky emigrant upozornil Spartanov, ze ich chce prepadnut perzsky kral Xerxes. Pouzil na to voskovane tabulky z dreva. Najprv vosk odstranil, vyril spravu na drevenu tabulku a znova tabulku zavoskoval. Tabulky prepasoval pomocou poslov. Greci najprv nevedeli, co to ma znamenat. Dostali napad: vosk nad ohnom roztavili a objavila sa sprava.

Takyto druh utajovania sprav sa vola **steganografia**. Utajuje sa nielen obsah spravy, ale aj

samotna sprava. Aj dnes sa pouziva "necitatelny" atrament alebo fotograficka "bodka". Existuju krajiny, kde sa nesmu posielat zakodovane e-maily.

2. Pre nas je typicky priklad zo zaciatku nasho letopoctu: Rimsky vojvodca a neskorsi cisar Cezar poslal zasifrovanu spravu jednej rimskej posadke obklucenej keltskymi vojskami v dnesnej Francii. V sprave ich vyziva, aby vydrzali, ze im ide na pomoc. Sprava bola dvojakym sposobom utajena. Pouziva grecku abecedu namiesto rimskej, a co je podstatne, slova utvoril posunutim pismen v ramci abecedy, t.j. namiesto pismena x napisal pismeno $x+3$. Napr. miesto slova "ano" by napisal slovo "dqr". Takyto kod sa zvykne volat Cezarov.

3. Arabi (8. - 10. stor.) prisli s napadom, ako desifrovat spravy. Zacali si vsimat pocet vyskytov jednotlivych pismen.

4. R. 1585 podstatne vylepsil Francuz Vigenere Cezarov kod. Objavil sa po prvykrat novy pojem ako "kluc" kodu. Vyse 200 rokov sa nedarilo najst vhodny algoritmus na desifrovanie tohto noveho kodu.

5. Po prvej svetovej vojne rapidny rozvoj sifrovacich strojov. Najzaujimavejsi je nemecky pristroj s menom Enigma. Okolo r. 1942 sa podarilo Anglicanom najst algoritmus na desifrovanie.

6. Enormny rozvoj kryptografie po druhej svetovej vojne. Hlavne s prichodom pocitacov.

Kodovanie tajnych sprav (2. lekcia)

Sifrovanie a **desifrovanie** budu nase nove pojmy. Budu sa castejsie pouzivat v zmysle **kodovanie** a **dekodovanie**. Dalej to bude **zdrojova abeceda**, v ktorej je napisana povodna sprava (angl. **plain-text**). Nasleduje **sifrovacia abeceda**, resp.

kodovacia abeceda, v ktorej je napisana zasifrovana sprava. Prichadza dolezita

Definicia. **Kryptosystem** volame paticu neprazdnych konecnych mnozin (P,C, K,E,D) s nasledovnymi vlastnostami:

1. P pozostava z textov (slov) nad zdrojovou abecedou.

2. C su texty (slova) nad sifrovacou abecedou.

3. K je mnozina **klucov**.

4. $E = \{E_k : k \in K\}$ je mnozina zobrazeni E_k : $P \rightarrow C$, ktore volame **sifrovacimi** funkciami.

5. $D = \{D_k : k \in K\}$ je mnozina **desifrovacich** funkcii $D_k : C \rightarrow P$.

6. Pre kazde $e \in K$ existuje take $d \in K$, ze $D_d(E_e(p)) = p$ pre vsetky $p \in P$.

Poznamka. Vsimnime si, ze sifrovacie funkcie su proste zobrazenia. V pripade $P{=}C$ su to uz bijekcie, prakticky akesi permutacie. Preco?

Pripominame, ze (skoro) vsetky mnoziny, o ktorych tu budeme hovorit su konecne a neprazdne. A^* znamena mnozinu vsetkych slov $a_1 a_2 \cdots a_k$, kde $a_i \in A$. Aj prazdne slovo tam patri. A volame pritom abecedou pre A^*. Dalej, A^n su slova nad A **dlzky** n.

Priklady na kryptosystem. V prvom pripade bude $P = C = K = \Sigma = \{A, B, \ldots, Z\}$, kde posledna mnozina predstavuje nasu beznu velku abecedu bez pismen s dlznami a hacikmi. Casto pouzijeme aj malu abecedu $\{a, b, \ldots, z\}$ ako P a $\Sigma = C$. Navyse, mame znamu bijekciu $A \mapsto 0$, $B \mapsto 1$, . . . , $X \mapsto 23$, $Y \mapsto 24$, $Z \mapsto 25$ medzi Σ

a (okruhom zvyskovych tried) Z_{26}. V dalsom budeme casto pracovat s okruhmi zvyskovych tried Z_n.

Priklad 1. Vezmime Σ a identifikujme tuto mnozinu s Z_{26}. Sifrovacie a desifrovacie funkcie budu $E_e : \Sigma \to \Sigma$ a $D_d : \Sigma \to \Sigma$, pricom

$x \mapsto (x + e)(\text{mod } 26)$ a $\text{y} \mapsto (\text{y} - \text{d})(\text{mod } 26)$.

Pri desifrovani potrebujeme polozit $d = e$.

Poznamka. Tento sposob sifrovania volame **posuvom**. Prvky mnoziny Z_n sa cyklicky posunuli o e miest. Je to vlastne specialna permutacia (substitucia) na Z_n. Pri $e = 3$ dostavame nasu znamu Cezarovu sifru (kod). Vo vseobecnosti mame dve velke moznosti, ako pozmenit text: **transpoziciou** alebo **substituciou** (permutaciou). Uvedieme jednoduchy priklad.

Priklad 2. Mame text: *Stretneme sa o pol noci*. Prepiseme ho tak, aby mal parny pocet

pismen a vsetky napiseme spolu: *stretnemesaopolnocii*. Urobime **transpoziciu**. Najprv napiseme v tom istom poradi pismena stojace na neparnom mieste a potom pridu pismena z parnych miest. Dostaneme zasifrovany text: *SRTEEAPLOITENMSOONCI*. Mozeme to este skomplikovat. Najprv napiseme pismena leziace na miestach, ktore po deleni cislom 3 daju zvysok 1. Potom nasleduju pismena napisane na miestach, ktore po deleni cislom 3 davaju zvysok 2. Nakoniec sa napisu pismena leziace na miestach delitelnych cislom 3. Vo vseobecnosti vidime, ze si mozeme vybrat cislo k a podla zvysku po deleni cislom k usporiadame pismena napisane v danom texte. V kazdom pripade sme urobili transpoziciu textu.

Priklad 3. Urobme na Z_{26} nasledovnu substituciu: *(a,X), (b,N), (c,Y), (d,A), (e,H), (f, P), (g,O), (h,G), (i,Z), (j,Q), (k,W), (l,B), (m,T), (n,S), (o,F), (p,L), (q,R), (r,C), (s,V), (t,M),*

(u,U), (v,E), (w,K), (x,J), (y,D), (z,I). Mame permutaciu π na nasej abecede, pricom $\pi(a) = X$, ..., $\pi(z) = I$. Zoberme text z prikladu 2 a pouzijme nan permutaciu π. Dostaneme zasifrovanu spravu: VMCHMSHTHVXFLFB-SFYZ. Desifruje sa to pomocou permutacie π^{-1}. Zrejme permutacie π a π^{-1} patria medzi kluce. Vidime, ze $\pi \neq \pi^{-1}$, pretoze $\pi(a) = X$, ale $\pi^{-1}(a) = D$.

Poznamka. Kryptosystem sa nazyva **symetrickym**, ak v podmienke 6 definicie kryptosystemu je vzdy $e = d$, alebo d sa da aspon lahko vypocitat s e. V opacnom pripade hovorime o **asymetrickom** systeme.

Blokove sifry (3. lekcia)

Definicia. **Blokova sifra** (kod) je kryptosystem, v ktorom P=C= Σ^n a K= $S(\Sigma^n)$ pre nejake prirodzene cislo n (Σ je lubovolna abeceda),

pricom $S(A)$ znamena mnozinu vsetkych permutacii (substitucii) na A. Najcastejsie sa pouziva blokova sifra, kde $K = S_n$.

Priklad 4. Blokova sifra sa vyznacuje sifrovacou funkciou

$$E_\pi : \Sigma^n \to \Sigma^n, \ \ (v_1, \cdots, v_n) \mapsto (v_{\pi(1)}, \cdots, v_{\pi(n)}).$$

Desifrovacia funkcia vyzera podobne, len substitucia je π^{-1}. Ak dodame, ze $P = C = \Sigma^n$ a $K = S_n$, tak dostavame tzv. **permutacnu sifru**.

Poznamka. Casto sa pouziva viacnasobne zasifrovanie. Jedna sa konkretnejsie o tzv. E-D-E trojnasobne sifrovanie. Potom to vyzera nasledovne:

$$c = E_k(D_l(E_m(p))).$$

D.U. Ako vyzera desifrovacia funkcia ku E-D-E trojnasobnemu sifrovaniu?

Veta 1. Nech $Z_n = (Z_n; \oplus, *)$ znamena okruh zvyskovych tried modulo n. Nech $0 \neq a \in Z_n$. Potom k prvku a existuje v Z_n inverzny prvok (t.j. prvok $b \in Z_n$ s vlastnostou $a * b = 1$) prave vtedy, ked $\gcd(a, n) = 1$ v Z.

Dokaz. Predpokladajme, ze existuje inverzny prvok $b = a^{-1}$. (Vieme, ze je len jediny!) Urobime nepriamy dokaz. Nech $\gcd(a, n) = d \neq 1$, t.j. cisla a a n su sudelitelne. Potom $a = d \cdot a_1$ a tiez $n = d \cdot n_1$ v Z. Zrejme $0 \neq n_1 \in Z_n$. Dalej,

$$a \cdot n_1 = (a_1 \cdot d)n_1 = a_1(d \cdot n_1) = a_1 \cdot n.$$

Vyslo nam, ze $a * n_1 = 0$ v Z_n. Dostavame

$$0 = b * 0 = b * (a * n_1) = (b * a) * n_1 = n_1,$$

co je spor, pretoze $n_1 \neq 0$.

Obratene, nech $\gcd(a, n) = 1$ v Z. Zo znamej vety o n.s.d. vieme, ze existuju cisla $u, v \in Z$ s vlastnostou

$$au + nv = 1.$$

Ak $0 \leq u < n$, tak je dobre. (Co ak $u = 0$?) Ak nie, tak vydelime (v okruhu Z) u s n a dostaneme $u = nq + b$, kde $0 \leq b < n$. Dalej,

$$1 = au + nv = a(nq + b) + nv =$$

$$ab + n(aq + v) \equiv ab(\text{mod n})$$

(Preco $b \neq 0$?) Inac povedane, dokazali sme $a * b = 1$, co bolo treba dokazat.

Definicia. Nech m je prirodzene cislo. Potom **afinnou sifrou** je blokova sifra, pricom $\Sigma = Z_m$ a dlzka bloku je $n = 1$. Dalej, mnozina klucov K pozostava z usporiadanych dvojic $(a, b) \in Z_m^2$, pricom a je nesudelitelne s cislom m. Sifrovacia funkcia E_e ku klucu $e = (a, b)$ je

$$E_e : \Sigma \to \Sigma, \quad x \mapsto ax + b(\text{mod m}).$$

Desifrovaciu funkciu ku klucu $d = (a', b)$ definujeme nasledovne:

$$D_d : \Sigma \to \Sigma, \quad x \mapsto a'(x - b)(\text{mod m}).$$

Cislo a' dostaneme ako riesenie rovnice $ax \equiv 1(\text{mod m})$ (alebo ako riesenie rovnice $a * x = 1$ v Z_m).

Priklad 5. Zoberme $\Sigma = Z_{26}$, t.j. "rimsku" abecedu. Vyberme napr. kluc $(a, b) = (7, 3)$. Chceme zasifrovat text COCA COLA, co prepiseme do jedneho slova: COCACOLA. Tomu odpoveda ciselna postupnost 2, 14, 2, 0, 2, 14, 11, 0. Tieto cisla zasifrujeme afinnym kodom a dostaneme novu postupnost cisel: 17, 23, 17, 3, 17, 23, 2, 3. Odpoveda to zasifrovanemu slovu: RXRDRXCD. Ukazte, ze a' v nasom pripade bude 15. Mozeme to pouzit na desifrovanie a presvedcime sa, ze sme dobre kodovali.

Linearne afinna sifra (4. lekcia)

Nech R je okruh s jednotkou. Hovorime, ze $u \in R$ je *delitelom jednotky*, ak ku u existuje inverzny prvok. Maticu A typu k/m nad R

budeme zapisovat ako $A = (a_{ij}) \in R^{(k,m)}$. Ak $A, B \in Z^{(n,n)}$, tak piseme $A \equiv B$(mod m) prave vtedy, ked $a_{ij} \equiv b_{ij}$(mod m).

Definicia. Zobrazenie $f : Z_m^{(n,1)} \to Z_m^{(n,1)}$ volame *linearne afinnym*, ak existuju matica $A \in Z_m^{(l,n)}$ a vektory $\mathbf{b} \in Z_m^{(l,1)}$ a $\mathbf{v} \in Z_m^{(n,1)}$ take, ze

$$f(\mathbf{v}) = A\mathbf{v} + \mathbf{b}.$$

Ak $\mathbf{b} = \mathbf{0}$, tak f sa vola linearnym zobrazenim.

Pripominame, ze riadkove vektory sa zapisuju po zlozkach nasledovne: $\mathbf{v} = (v_1, \cdots, v_n) \in Z_m^{(1,n)}$, co podla nasho dohovoru znacime ako Z_m^n. Stlpcove vektory su z mnoziny $Z_m^{(n,1)}$.

Definicia. Blokovu sifru s dlzkou bloku n nazyvame **linearne afinnou sifrou**, ak $P = C = Z_m$ a vsetky sifrovacie funkcie su linearne afinne. Specialne, sifra sa vola **linearnou**, ak sifrovacie funkcie su linearne.

Rozpiseme definiciu podrobnejsie. Potrebujeme vediet, co su kluce a ako vyzeraju sifrovacie a desifrovacie funkcie. Zrejme, sifrovacie funkcie maju tvar (ak stotoznime Z_m^n s $Z_m^{(n,1)}$)

$$E : Z_m^n \to Z_m^n, \quad \mathbf{v} \mapsto A\mathbf{v} + \mathbf{b},$$

kde E by malo byt prostym zobrazenim. Nas pripad si zjednodusime, ak budeme predpokladat, ze $l = n$, t.j. A bude stvorcovou maticou stupna n. Z linearnej algebry vieme (?), ze to nastane prave vtedy, ked $\det(A)$ je delitelom jednotky v Z_m. Teda, klucom je dvojica $(A, \mathbf{b})$, kde A je stvorcova matica stupna n nad Z_m a $\mathbf{b}$ stlpcovy vektor dlzky n a plati, ze $(\det(A), m) = 1$ v Z (vid nasu predchadzajuvcu vetu). Potom (znova z linearnej algebry) mame

$$A^{-1} = (\det(A))^{-1} \cdot \mathrm{adj}(A),$$

kde $\mathrm{adj}(A)$ je adjungovana matica ku matici A. (Ostava zaujimavy pripad $l \neq n$. Tu sa hladaju podmienky na existenciu lavej inverznej matice

B typu n/l ku A, t.j. $B \cdot A = E_n$, kde teraz E_n znamena jednotkovu maticu stupna n.)

Desifrovacie funkcie sa vypocitaju teraz jednoducho (Overte to!):

$$D : Z_m^n \to Z_m^n, \quad \mathbf{v} \mapsto A^{-1}(\mathbf{v} - \mathbf{b}).$$

Uvedieme dva typicke priklady na afinne linearne sifry.

Priklad 6. Jedna sa o Vigenereovu sifru. (Blaise Vigenere zil v 16. storoci.) Mnozina klucov je dana vztahom $K = Z_m^n$. Ak $\mathbf{k} \in Z_m^n$, tak mame

$$E_{\mathbf{k}} : Z_m^n \to Z_m^n, \quad \mathbf{v} \mapsto \mathbf{v} + \mathbf{k} \quad a$$

$$D_{\mathbf{k}} : Z_m^n \to Z_m^n, \quad \mathbf{v} \mapsto \mathbf{v} - \mathbf{k}.$$

Zrejme $|K| = m^n$.

Priklad 7. Pojde o Hillovu sifru. (Lester S. Hill ju vynasiel r. 1929.) Mnozina klucov K

pozostava zo vsetkych stvorcovych matic A stupna n nad Z_m, pricom ziadame, aby platilo $(\det(A), m) = 1$ v Z. Potom pre $A \in K$ mame

$$E_A : Z_m^n \to Z_m^n, \quad \mathbf{v} \mapsto A\mathbf{v}.$$

Desifrovacie funkcie D_A sa ziskaju uz jednoducho. Hillova sifra je linearnou sifrou.

Priklad 8. Vratme sa ku permutacnej sifre (priklad 4). Predpokladajme vsak, ze $\Sigma = Z_m$. Ukazeme, ze nasa permutacna sifra je specialnym pripadom linearnej sifry, ba dokonca Hillovej sifry. Nech E_n znamena teraz jednotkovu maticu stupna n. Z linearnej algebry vieme, ze vymena i-teho riadku s j-tym v matici A sa da vykonat aj tak, ze vynasobim $E_{ij} \cdot A$, pricom E_{ij} je matica, ktora vznikne z E_n, ak na nej vymenim i-ty riadok s j-tym. Dalej vieme, ze kazda permutacia $\pi \in S_n$ sa da napisat ako sucin vhodnych transpozicii. Dalej si vsimnime, ze

$$(v_1, \cdots, v_{i-1}, v_j, \cdots, v_i, v_{j+1}, \cdots, v_n) = E_{ij} \cdot \mathbf{v}.$$

Transpozicia (i, j) odpoveda elementarnej matici $E_{i,j}$. Nech teraz π je permutacia z S_n. Vykonajme v sulade s π vymenu riadkov na jednotkovej matici E_n. Dostaneme maticu E_π. Vzhladom na to, ze π je sucinom vhodnych transpozicii $\pi = (l, k) \cdots (i, j)$ v S_n, tak pre matice plati

$$E_\pi = E_{l,k} \cdots E_{i,j} \cdot E_n.$$

Preto mame

$$(v_{\pi(1)}, \cdots, v_{\pi(n)}) = E_\pi \cdot \mathbf{v},$$

co je Hillova sifra.

Kryptoanalyza (5. lekcia)

Spomenieme niekolko technik, ktore sa pouzivaju pri "utokoch" (alebo pri rozuzleni) na zakodovane texty. Najprv urobime vseobecny predpoklad, tzv. Kerckhoffov princip: Zhruba poznáme pouzity kryptosystem. Co nepoznáme, je kluc a povodny text (plaintext).

Existuje viacej urovni utokov na kryptosystemy. Rozdelujeme ich nasledovne:

1. **Ciphertext-only**. Pozname retazec sifrovaneho textu: **y**. K tomu hladame odpovedajuci povodny text **x** a kluc.

2. **Known plaintext**. Pozname retazec povodneho textu **x** a odpovedajuci retazec sifrovaneho textu **y**. Hladame kluc a pokusame sa desifrovat dalsie casti textu.

3. **Chosen plaintext**. Mali sme urcitu dobu pristup k sifrovaciemu zariadeniu. Vieme vyrobit sifrovane texty **y** ku vybranym povodnym textom **x**. Hlada sa kluc.

4. **Chosen ciphertext**. Mali sme urcitu dobu pristup ku desifrovaciemu zariadeniu. Ku retazcu zasifrovaneho textu **y** vieme vyrobit patricny retazec povodneho textu **x**. Hladame kluc.

Zaciname obycajne s najslabsim utokom, t.j. ciphertext only. Samozrejme, predpokladame (vieme), ze hladany text je v slovencine (anglictine a pod.). Text je napisany bez interpunkcii a medzier. (To este viacej stazuje desifrovanie.) Utok pozostava z toho, ze pomocou pocitaca sa snazime postupne pouzivat vsetky kluce a desifrovat zasifrovany text. Ten spravny text by mal byt medzi nimi.

Pri utoku pomocou known plaintextu casto vyuzivame statisticke vlastnosti daneho jazyka (slovenciny, anglictiny). Totiz, ak pouzivame kryptosystem odpovedajuci posunu pismen (Cezar), tak pri pevnom kluci, pismeno v povodnom texte sa rovnako casto objavi ako jemu odpovedajuca sifra. (Prisli na to uz Arabi!) Napr. v anglictine je pravdepodobnost vyskytu pismena *E* 0,120. Pravdepodobnost vyskytu pismen *T, A, O, I, N, S, H, R* sa pohybuje medzi 0,06 az 0,09. Pismena *D, l* maju pravdepodobnost vyskytu okolo 0,04. Dalej *C, U, M, W,*

F, G, Y, P, B maju pravdepodobnosti vyskytu medzi 0,015 a 0,028. Nakoniec, *V, K, J, X, Q, Z* maju pravdepodobnosti vyskytu mensie ako 0,01.

Podobne ako vyssie, oplati sa si vsimat dvojice, resp. trojice za sebou iducich pismen, tzv. *digramy* resp. *trigramy*. Najcastejsie sa vyskytujuci digram v zasifrovanom texte odpoveda najcastejsie sa vyskytujucemu digramu v povodnom texte, teda napr. v anglictine. V anglictine je znamych 30 castych digramov. Su to (usporiadane podla klesajuceho vyskytu): *TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, AR, TE, SE, HI, OF*. Podobne mame 12 najcastejsich trigramov *THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH*.

Kryptoanalyza linearne afinnej sifry

Predpokladame, ze mame do cinenia z linearne afinnou sifrou nad abecedou Z_m s dlzkou bloku n. Mame nejaky kluc. Potom sifrovacia funkcia je tvaru

$$E : Z_m^n \to Z_m^n, \quad \mathbf{v} \mapsto A\mathbf{v} + \mathbf{b} \text{ (mod m)},$$

kde $A \in Z^{(n,n)}$ a $\mathbf{b} \in Z^n$. Chceme zistit kluc $(A, \mathbf{b})$. K tomu pouzijeme $n+1$ blokov "povodneho" textu $\mathbf{w}_i$ pre $0 \leq i \leq n$. Tomu odpoveda sifrovany text (pri danom kluci) tvaru

$$\mathbf{c}_i = A\mathbf{w}_i + \mathbf{b}, \quad 0 \leq i \leq n.$$

Potom

$$\mathbf{c}_i - \mathbf{c}_0 \equiv A(\mathbf{w}_i - \mathbf{w}_0)(\text{mod m}).$$

Nech W je matica

$$W = (\mathbf{w}_1 - \mathbf{w}_0, \cdots, \mathbf{w}_n - \mathbf{w}_0)(\text{mod m}),$$

kde $(\mathbf{w}_i - \mathbf{w}_0)(\text{mod m})$ pre $1 \leq i \leq n$ tvoria stlpce matice W. Podobne C je matica

$$C = (\mathbf{c}_1 - \mathbf{c}_0, \cdots, \mathbf{c}_n - \mathbf{c}_0)(\text{mod m}).$$

Dostaneme,

$$AW \equiv C(\text{mod m}),$$

co sa da prepisat na rovnost $A_1W_1 = C_1$ nad Z_m, kde A_1, W_1 a C_1 su matice nad Z_m a plati $A \equiv A_1(\text{mod m})$, $W \equiv W_1(\text{mod m})$ a $C \equiv C_1(\text{mod m})$. Podla vety 1 potrebujeme zistit, ci $(\det(W_1), m) = 1$, alebo nie. V prvom pripade existuje inverzna matica W^{-1} ku W nad Z_m. Dostaneme

$$A \equiv C(w'adj(W))(\text{mod m}),$$

pricom w' znamena inverzny prvok ku $\det(W_1)$ nad Z_m. Takto mozeme vypocitat

$$\mathbf{a}_1, \cdots, \mathbf{a}_n,$$

stlpce matice A. Nakoniec ziskame

$$\mathbf{b} = \mathbf{c}_0 - A\mathbf{w}_0$$

a mame kluc. Ak je sifra linearna, tak staci polozit $\mathbf{w}_0 = \mathbf{c}_0 = \mathbf{0}$. Dostaneme $\mathbf{b} = \mathbf{0}$.

Perfektne sifry (6. lekcia)

Po vsetkych doterajsich skusenostiach sa pytame, ci existuju sifry (kody), ktore by boli stopercentne spolahlive pred neziaducim desifrovanim, t.j. ktore by sa nedali "zlomit". Odpoved znie ANO, ovsem za aku cenu?

Samozrejme, potrebovali by sme najprv matematicku definiciu perfektnej sifry. Nebudeme sa do toho pustat, ale upozornujeme, ze existuje teoria opisana Claude Shannonom v r. 1949, ktora sa takymi kryptosystemami zaobera. Opiseme si jeden taky kryptosystem. Autorom je Gilbert Vernam, ktory predlozil r. 1917 navrh noveho kryptosystemu a nechal ho aj patentovat. (Vtedy slo o telegraficke spravy.) Opiseme tento **Vernam-One-time-Pad** kryptosystem.

Nech $n \geq 1$ je prirodzene cislo. Polozme

$$P = C = K = (Z_2)^n = Z_2^n.$$

Ak $x = (x_1, \cdots, x_n) \in P$ a $K = (K_1, \cdots, K_n) \in K$, tak sifrovacie a desifrovacie funkcie su

$$e_K(x) = (x_1 \oplus K_1, \cdots, x_n \oplus K_n),$$

$$d_K(y) = (y_1 \oplus K_1, \cdots, y_n \oplus K_n),$$

pre $y = (y_1, \cdots, y_n) \in C$, kde $(Z_2; \oplus, *, 0, 1)$ je okruh zvyskovych tried modulo 2.

Uvedieme jednoduchy priklad. Najprv zasifrujeme abecedu blokovym kodom nad Z_2. Dlzka bloku je $m = 5$ a definujeme nasledovne priradenie: a=00000, b=00001, c=00010, d=00011, e=00100, f=00101, g=00110, h=00111, i=01000, j=01001, k=01010, l=01011, m=01100, n=01101, o=01110,p=01111,q=10000, r=10001, s=10010, t=10011, u=10100, v=10101, w=10110, x=10111, y=11000, z=11001. Chceme poslat (binarny) text

$$01010100011100001111001101110$$

$$00110100010000000101010000000,$$

v ktorom je 60 binarnych znakov. Kluc (”dazdovka”) by mal mat tiez aspon 60 binarnych znakov. Mali by sa vybrat naprosto nahodne. Zvolme teda

100010011010111100011011000101

011001010010101000111010101110.

Zasifrujeme nas text podla Vernama, t.j. scitame nas (binarny) text s klucom (dazdovkou) modulo 2. Dostaneme zasifrovany text

110111011101111111100010101011

010100010110101001101110101110.

Ten teraz desifrujeme, ak pripocitame znova nas kluc (dazdovku) ku ziskanemu textu. (Nezabudajme, ze kluc pozna aj prijimatel.) Ak sme sa nepomylili, tak dostaneme povodny binarny text a ten este prelozime do povodneho jazyka. (Riesenie: kryptografia.)

Existuju rozne obmeny Vernamovho kryptosystemu. Napr. namiesto Z_2 vezmeme Z_{10}. Ako priklad mozeme spomenut kod, ktory pouzival Che Guevara pri tajnom kontakte s Fidelom Castrom. Naslo sa to pri mrtvom tele Ch. Guevaru r. 1967 v Bolivii. Beznu abecedu mal zakodovanu v decimalnom systeme: a=6, b=38, c=32, d=4, e=8, f=30, g=36, h=34, i=39, j=31, k=78, l=72, m=70, n=76, o=9, p=79, q=71, r=58, s=2, t=0, u=52, v=50, w=56, x=54, y=1, z=59. (Takato volba zodpoveda istym specifikam spanielskeho jazyka.) Teraz pouzil Vernamov kryptosystem nad Z_{10}, t.j. scitoval modulo 10. (Postupnost cisel mal usporiadanu po paticiach.) Zakodovany text posielal pomocou vysielacky. (Kluc poznal len on a Castro.) Podobne aj on prijimal spravy z Kuby tymto sposobom.

Pseudonahodne postupnosti

Vernamov kryptosystem sa udajne pouziva pri tzv. horucom telefone spajajucom Washington s Moskvou.

Hoci je Vernamov system naprosto spolahlivy, ma vela nevyhod a nepouziva sa masove. Velkou nevyhodou je, ze kluc musi byt aspon tak velky ako je vysielany text. Dalej tento kluc musi sa zvolit naprosto "nahodne", t.j. postupnost binarnych cifier musi sa vybrat "nahodne". Inac by sa naskytla moznost desifrovania (rozlustenia) zatajeneho textu. Neexistuje presna matematicka definicia toho, co je **nahodna** velicina, ci nahodna postupnost. Napr. mozno vybrat taku postupnost v priebehu nejakych merani vo fyzike jadra, kde vieme, ze plati nahoda. (T.j. nevieme ani teoreticky vyjadrit suvis hodnoty x_n s hodnotou x_{n+1} vzniknutej pri dalsom merani.) Namiesto toho pouzivame tzv. **pseudonahodne** cisla, ci postupnosti. Znama je Knuthova definicia pseudonahodnych postupnosti. Na vyrobu takych postupnsti mame

generatory nahodnych velicin. Nakoniec, pseudonahodnu postupnost (kluc) musime tajne dorucit partnerovi, inac cela procedura utajovania nema zmysel. Tieto nevyhody zapricinili, ze Vernamov kod pouzivaju len diplomati, spioni, banky a milenci.

Enigma

Do druhej svetovej vojny sa sifrovanie (kodovanie) a desifrovanie vykonavalo rucne a zdlhavym sposobom. Az pocas vojny sa na tuto pracu zaviedli elektronicko-mechanicke pristroje. Najznamejsi z nich je nemecka Enigma. Americania mali svoj sifrovaci stroj M-209. Podobne mali Anglicania, ci Japonci tiez svoje sifrovacie a desifrovacie pristroje.

Nemci boli az do konca vojny presvedceni, ze ich Enigma je neprekonatelna. Ukazalo sa, ze spojenci dokazali desifrovat nemecke sifrovane

spravy uz v r. 1942. To malo urcite aj vplyv na priebeh vojny.

DES-algoritmus (7.lekcia)

Po druhej svetovej vojne nastal v USA rychly rozvoj pocitacovej techniky. S prichodom tranzistorov (1949) a integrovanych obvodov (microchips) v r. 1959 sa stali pocitace dostupne aj pre sirsiu verejnost. Okrem armady a vladnych uradov zacali pocitace pouzivat aj podniky a banky. Pri transfere penazi a obchodnych tajomstiev sa kladol velky doraz na utajovanie udajov. K slovu prisla kryptografia, ovsem s novymi problemami.

Sifrovanie sprav pomocou pocitacov sa v podstate podoba staremu procesu, ked sa pracovalo rucne alebo s mechanizmami. Su tu vsak tri rozdiely: 1. pomocou pocitaca sa daju riesit

omnoho komplikovanejsie ulohy ako s mechanickymi sifrovacimi strojmi, 2. pocitace pracuju omnoho rychlejsie a 3. pocitace pracuju len s binarnymi cislami.

R. 1973 sa vypisal konkurz na projekt standartizovaneho kryptosystemu, alebo prakticky pouzitelneho sifrovacieho algoritmu, ktory by umoznoval podnikom spolahlivo, rychlo a tajne komunikovat s inymi podnikmi. Konkurz vyhral produkt firmy IBM s krycim menom Lucifer. Vynasiel ho isty Horst Feistel. Dnes to poznáme pod menom **DES-algoritmus**, pricom DES znamena *Data Encryption Standard*. (Oficialne sa to zaviedlo koncom roku 1976.) DES-algoritmus je specialnym pripadom tzv. Feistelovej sifry (sietoveho systemu).

Prv nez vysvetlime Feistelove sifry, tak si pripomenieme, ako sa prepise (zakoduje) bezna abeceda pomocou postupnosti zlozenych z 0 a 1. Jedna

sa o americku normu **ASCII** (*American Standard Code for Information Interchange*). Na kazdy znak pouzijeme 7-miestne binarne cislo. Takto mozeme zakodovat $128 = 2^7$ pismen a znakov. Napr. pismeno a=1100001, alebo !=0100001. Kvoli lepsej predstavivosti si uvedieme ASCII-kody velkych pismen: A=1000001, B=1000010, C=1000011, D=1000100, E=1000101, F=1000110, G=1000111, H=1001000, I=1001001, J=1001010, K=1001011, L=1001100, M=1001101, N=1001110, O=1001111, P=1010000, Q=1010001, R=1010010, S=1010011, T=1010100, U=1010101, V=1010110, W=1010111, X=1011000, Y=1011001, Z=1011010. V praxi pouzivame vsak kvoli lepsej kontrole 8-ice, kde poslednu, t.j. 8. zlozku, doplnime 0 alebo 1 na zaklade celkovej parity (aby sme mali parny pocet 1-ciek).

Opiseme Feistelov sietovy system (na vseobecnej urovni). Budeme pracovat so slovami dlzky

$2t$ nad abecedou $Z_2 = \{0, 1\}$. Slova dlzky $2t$ mozeme zapisovat ako

$$p = (L_0(p), R_0(p)),$$

kde $L_0(p)$ a $R_0(p)$ su podslova dlzky t. ($L_0(p)$ volame lavou polovickou slova p a $R_0(p)$ zase pravou polovickou.) Slovo p je zlozenim slov $L_0(p)$ a $R_0(p)$, t.j. $p = L_0(p)R_0(p)$. Teraz pojde o tri dolezite kroky:

1. Zvolili sme si permutaciu, tzv. **inicialnu**, $IP \in S_{2t}$. Tuto pouzijeme na slovo p, t.j. urobime $IP(p)$. Polozme $IP(p) = (L_0, R_0)$.

2. (Najdolezitejsi krok): Urobime $r \geq 1$ iteracii (opakovani) (r je dopredu zadane cislo),

$$((L_i, R_i) : 0 \leq i \leq r)$$

podla nasledovneho receptu

$$(L_i, R_i) = (R_{i-1}, L_{i-1} \oplus f_{K_i}(R_{i-1})),$$

kde treba este vysvetlit, co znamenaju K_i a ako pracuje funkcia f_{K_i}. Nakoniec,

3. Pouzijeme inverznu permutaciu $(IP)^{-1}$ na slovo R_rL_r, t.j. $(IP)^{-1}(R_rL_r)$ a to je zaver vseobecnej casti. Totiz,

$$E_k(L_0, R_0) = (IP)^{-1}(R_rL_r)$$

je sifrovacia funkcia. Desifrovaciu funkciu D_k ziskame pomerne lahko. Postupujeme opacnym smerom ako pri sifrovani. Opiseme to strucne. Nech $y = y_1y_2 \cdots y_s$ je tajna sprava, ktoru sme ziskali. Podslova $y_1, \cdots, y_s$ su dlzky $2t$ a len s nimi sa budeme zaoberat, lebo vieme, ze sa jedna o DES-kryptosystem. Vyberme si y_j. Najprv permutujeme pomocou IP, t.j.

$$(IP)(y_j) = R_rL_r = (R_r, L_r).$$

Na zaklade bodu 2 dostavame $R_{r-1} = L_r$ a sucasne

$$L_{r-1} = R_r \oplus f_{K_r}(L_r),$$

lebo pracujeme nad Z_2. Patricne vela razy to zopakujeme, az dostaneme slovo $x = (L_0, R_0)$ dlzky $2t$. Urobime este $(IP)^{-1}(x)$ a dostaneme povodnu spravu napisanu v binarnej abecede.

V dalsej casti podrobnejsie opiseme prislusny kryptosystem (P,C,K,E,D) a pritom aj K_i a f_{K_i}.

DES-algoritmus (Pokracovanie) (8. lekcia)

Ak vo Feistelovom sietovom systeme upresnime potrebne udaje a dodame niektore nove udaje, tak sa dostaneme ku prakticky pouzitelnemu DES-algoritmu, ktory sa v USA pouziva este aj dnes. (Medzicasom bol vypisany konkurz na novy system.) Polozime $P=C=(Z_2)^{64}$, t.j. $t = 32$. Pracujeme so 64-ticami zlozenych z 0 a 1, ktore volame "bitami". Takyto vektor pozostava vlastne z 8 usporiadanych osmic, ktore volame "bajtami". Urcime si mnozinu klucov

$$K = \{(b_1, \cdots, b_{64}) \in (Z_2)^{64} : \sum (b_{8j+i} \equiv 1 (\mathrm{mod}\ 2))\}$$

pricom scitavame pre $1 \leq i \leq 8$ pri pevnom j. Parameter j je urceny vztahom $0 \leq j \leq 7$. Z toho vidime, ze mame len 56 volitelnych "bitov". To dava 2^{56} klucov. Ako priklad mozeme uviest jeden kluc:

00010011001101000101011101111001

10011011101111001101111111110001.

Dalej, pre nas system plati $r = 16$. Uvedieme tiez priklad takej permutacie $IP \in S_{64}$, ktora sa pouziva v praxi:

58 50 42 34 26 18 10 2 60 52 44 36 28 20 12 4

62 54 46 38 30 22 14 4 64 56 48 40 32 24 16 8

57 49 41 33 25 17 9 1 59 51 43 35 27 19 11 3

61 53 45 37 29 21 13 5 63 55 47 39 31 23 15 7.

Ku klucu $k \in K$ skonstruujeme $(K_1, \cdots, K_{16})$, co budeme potrebovat ku definicii f_{K_i}.
Zrejme,

$$k = k_1 \cdots k_{64} = k_1^* \cdots k_8^*,$$

kde k_i^* je 8-tica binarnych cislic, t.j. byte. Pretoze $k \in K$, tak mozeme z kazdeho k_i^* vynechat poslednu cislicu. Dostaneme 7-tice k_i'. Povodny kluc k sa zmenil na binarnu 56-ticu $k' = k_1' \cdots k_8'$. Slovo k' spracujeme pomocou dvoch permutacii (injektivnych zobrazeni) $PC-1$ a $PC-2$.
$PC-1 \in S_{56}$ vyzera takto:

57 49 41 33 25 17 9 1 58 50 42 34 26 18

10 2 59 51 43 35 27 19 11 3 60 52 44 36

63 55 47 39 31 23 15 7 62 54 46 38 30 22

14 6 61 53 45 37 29 21 13 5 28 20 12 4.

Napr. $(PC-1)(k_1') = k_{57}k_{49}\cdots k_9$ alebo $(PC-1)(k_8') = k_{21}k_{13}\cdots k_4$.

$PC-2 \in S_{48}$ sa definuje zase nasledovne:

14 17 11 24 1 5 3 28 15 6 21 10

23 19 12 4 26 8 16 7 27 20 13 2

41 52 31 37 47 55 30 40 51 45 33 48

44 49 39 56 34 53 46 42 50 36 29 32.

Vratme sa ku k'. Dalsi postup vykoname v dvoch krokoch:

1. Polozme $(PC-1)(k') = (C_0, D_0) = C_0D_0$, kde C_0 ma 28 bitov a podobne aj D_0.

2. Predpokladajme, ze $1 \leq i \leq 16$. Pocitajme

$$C_i = LS_i(C_{i-1}), \quad D_i = LS_i(D_{i-1}),$$

kde LS_i znamena, ze na slove C_{i-1}, resp. na D_{i-1}, sa urobi cyklicky posun dolava o 1 alebo o 2 miesta. Zalezi to od indexu i. Ak $i = 1, 2, 9, 16$, tak sa posun vykona o jedno miesto, inac o dve. Potom dostaneme

$$K_i = (PC-2)(C_iD_i).$$

Napr. $(PC-2)(b_1 \cdots b_{56}) = b_{14}b_{17} \cdots b_{32}$.

Treba este urcit hodnoty $f_{K_i}(R_{i-1})$. To sa tiez prevedie vo viacerych krokoch. Zrejme $f_{K_i} : (Z_2)^{32} \rightarrow (Z_2)^{32}$ a $K_i \in (Z_2)^{48}$, t.j. R_{i-1} je binarne slovo dlzky 32. Kvoli zapisu oznacme K_i ako J a R_{i-1} ako A. Ide teda o vypocet $f_J(A)$. Urobi sa to nasledovne:

1. Slovo A sa rozsiri na slovo dlzky 48 pomocou pevne urcenej "expansion" funkcie E. Nasledujuca tabulka urcuje funckiu E.

32 1 2 3 4 5 4 5 6 7 8 9

8 9 10 11 12 13 12 13 14 15 16 17

16 17 18 19 20 21 20 21 22 23 24 25

24 25 26 27 28 29 28 29 30 31 32 1.

2. Vypocitajme $E(A) \oplus J = B$, kde $B = B_1B_2B_3B_4B_5B_6B_7B_8$ a B_j su 6-bitove slova.

3. V nasledujucom kroku pouzijeme S-boxy $S_1, \cdots, S_8$. Kazdy S_i je pevna matica typu 4×16, ktorej prvky su cisla $0, \cdots, 15$. Ak zoberieme slovo $B_j = b_1 b_2 b_3 b_4 b_5 b_6$ z predchadzajuceho bodu, tak $S_j(B_j)$ sa vypocita nasledovne: bity $b_1 b_6$, nam urcia binarny zapis cisla r oznacujuceho riadok matice S_j $(0 \leq r \leq 3)$. Zvysne 4 bity $b_2 b_3 b_4 b_5$ urcia zase binarny zapis cisla s stlpca matice S_j $(0 \leq s \leq 15)$. Potom

$$S_j(B_j) = S_j(r, s) = C_j = c_{j1} c_{j2} c_{j3} c_{j4}$$

je cislo z matice S_j na pozicii (r, s) napisane v binarnom zapise pre kazde $1 \leq j \leq 8$. Nakoniec

4. $C = C_1 C_2 C_3 C_4 C_5 C_6 C_6 C_7 C_8$ pozostava z 32 bitov. Slovo C permutujeme podla zadanej permutacie P. Vysledok: $P(C) = f_J(A)$. Permutacia P je definovana nasledovne:

16 7 20 21 29 12 28 17 1 15 23 26

5 18 31 10 2 8 24 14 32 27 3 9

19 13 30 6 22 11 4 25.

Treba poznamenat, ze hlavnym problemom DES-algoritmu je dorucenie kluca od odosielatela ku prijimatelovi a jeho dalsie utajnenie. DES je symetricky kryptosystem. Zvykne sa tiez pouzivat pojem *private-key system*, na rozdiel od *public-key* systemu, o ktorom bude rec neskorsie.

Prvocisla (9. lekcia)

Prv nez budeme hovorit o dalsich kryptosystemoch, doplnime si niektore poznatky z matematiky. Casto potrebujeme navrhnut lubovolne a hodne velke prvocisla. Ako sa to da urobit?

Veta 2. Nech n je zlozene prirodzene cislo. Potom existuje prvocislo p tak, ze $p \mid n$ a sucasne $p \leq \sqrt{n}$.

Dokaz. Na zaklade predpokladu existuju prirodzene cisla a a b tak, ze $n = ab$, kde $a > 1$ a $b > 1$. (Okruh Z je usporiadany obor integrity! Podobne pole Q, ci R!) Preto plati: bud $a \leq \sqrt{n}$, alebo $b \leq \sqrt{n}$. V opacnom pripade by bolo $a > \sqrt{n}$ a $b > \sqrt{n}$, co by implikovalo

$$n = ab > \sqrt{n} \cdot b > \sqrt{n}\sqrt{n} = n,$$

a to je uz spor. Nech teda $a \leq \sqrt{n}$. Kedze $a \neq 1$, tak existuje prvocislo $p \mid a$, ktore vlastne hladame.

Priklad. Je cislo $n = 15413 \in Z$ prvocislom? Da sa overit, ze $\lfloor \sqrt{n} \rfloor = 124$. Podla vety potrebujeme overit vsetky neparne prvocisla $p \leq 124$. Jedna sa o nasledovne prvocisla: 3, 5, 7, 11, 13, 17, 19, 23, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113. Pomocou kalkulacky zistime, ze ziadne z uvedenych prvocisel nedeli cislo n. Teda, dane cislo n je prvocislo.

Veta 3. (Mala Fermatova veta). Nech n je prvocislo. Potom v okruhu $(Z_n; +, *)$ plati: $a^{*(n-1)} = 1$ pre kazde $0 \neq a \in Z_n$.

Dokaz. Vieme, ze $(Z_n; +, *)$ je pole. (Preco?) Potom $Z_n^* = (Z_n - \{0\}; *)$ je komutativna grupa. Nech $0 \neq a \in Z_n$ je lubovolny prvok. Utvorme cyklicku podgrupu $[a]$ konecnej grupy $(Z_n^*; *)$. Vieme, ze

$$[a] = \{a^0, a^1, \cdots, a^{*(k-1)}\},$$

kde k je najmensie prirodzene cislo s vlastnostou $a^{*k} = a * \cdots * a = 1$ v nasej grupe Z_{n-1}. (Take k existuje, pretoze Z_n je konecne.) Tiez vieme, ze $[a]$ ma k prvkov. Podla Lagrangeovej vety dostavame $k \mid n - 1$. Z toho vyplyva $n - 1 = k \cdot t$ pre nejake prirodzene t a nakoniec $a^{*(n-1)} = (a^{*k})^{*t} = 1$.

Existuje ekvivalentna formulacia predchadzajucej Malej Fermatovej vety: Nech n je prvocislo. Potom plati $a^{n-1} \equiv 1 (\text{mod n})$ pre kazde $a \in Z$ s vlastnostou: $\gcd(a, n) = 1$ v Z.

Priklad. Vezmime cislo $n = 341 = 11 \cdot 31$. Da sa overit, ze plati

$$2^{340} \equiv 1(\text{mod } 341).$$

Naproti tomu

$$3^{340} \equiv 56(\text{mod } 341).$$

Poucenie: Mala Fermatova veta sa da tiez pouzit na zistovanie toho, ci nejake prirodzene cislo je zlozene, alebo nie. V nasom pripade vychadza, ze 341 je zlozene cislo (hoci sme to vedeli aj zo zadania).

Otazka: Plati, ze n je prvocislo prave vtedy, ked $a^{n-1} \equiv 1(\text{mod n})$ pre kazde $\gcd(a, n) = 1$?
Odpoved: Nie. Ukazu to dalsie uvahy.

Definicia. Nech n je neparne zlozene prirodzene cislo. Nech pre cele cislo a plati

$$a^{n-1} \equiv 1(\text{mod n}).$$

Potom n volame *pseudoprvocislom* pri baze a. Dalej, ak n je pseudoprvocislom pri vsetkych bazach a, pre ktore je $\gcd(a, n) = 1$, tak n sa vola *Carmichaelovym cislom*.

Je zname, ze najmensim Carmichaelovym cislom je $561 = 3 \cdot 11 \cdot 17$. Potrebujeme este jednu definiciu: Nech

$$n = p_1^{e_1} \cdots p_k^{e_k}$$

je kanonicky rozklad prirodzeneho cisla. Hovorime, ze n ma *jednoduche faktory,* ak $e_1 = \cdots = e_k = 1$.

Veta 4. Nech n je neparne zlozene prirodzene cislo. Potom n je Carmichaelovym cislom prave vtedy, ked

(i) n ma jednoduche faktory a

(ii) $p \mid n$ implikuje $p - 1 \mid n - 1$ pre lubovolne prvocislo p.

Dokaz vynechame.

Na zaklade vety 4 sa mozeme rychle presvedcit, ze hore uvedene cislo 561 je Carmichaelovym cislom. Preco nie je 33 Carmichaelovym cislom?

Millerov-Rabinov test

Nasledujuca veta (Millerov-Rabinov test) je istym prehlbenim Malej Fermatovej vety. Predpokladajme, ze $n \in N$ je neparne. Dalej, nech

$$s = \max\{r \in N : 2^r \mid n - 1\}.$$

Zrejme $s \geq 1$ je najvacsim exponentom s vlastnostou

$$n - 1 = 2^s \cdot d.$$

Za tychto predpokladov mozeme vyslovit

Veta 5. Nech n je neparne prvocislo a nech $a \in N$ je take, ze $\gcd(a, n) = 1$. Potom plati (vzhladom na horeuvedene oznacenie) bud

$$a^d \equiv 1 (\mathrm{mod}\ n),$$

alebo existuje take $r \in \{0, 1, \cdots, s-1\}$, ze plati

$$a^{2^r d} \equiv -1 (\mathrm{mod}\ n).$$

Dokaz. Nech $a \in N$ je take, ze $\gcd(a, n) = 1$. Specialne, $a \in \{1, \cdots, n-1\}$, lebo n je prvocislo. Nech $Z_n^* = \{1, 2, \cdots, n-1\} \subseteq Z_n$ znaci podgrupu multiplikativnej pologrupy $(Z_n; *)$ (vid Vetu 1). Potom bud $a^d \equiv 1 (\mathrm{mod}\ n)$, t.j. $a^{*d} = 1$ v Z_n^*, alebo $a^{*d} \neq 1$. Predpokladajme v dalsom druhy pripad.

Lenze a^{*d} generuje v Z_n^* netrivialnu cyklicku podgrupu $[a^{*d}]$. Vieme, ze

$$[a^{*d}] = \{1, a^{*d}, \cdots, a^{*(k-1)d}\},$$

kde k je najmensie cislo s vlastnostou $a^{*kd} = 1$ v Z_n^*. Podla Lagrangeovej vety mame

$$kd \mid n - 1 = 2^s \cdot d.$$

Pretoze Z je obor integrity, tak mozeme kratit s cislom d a dostavame $k \mid 2^s$ v N. Teraz uz vidime, ze $k = 2^t$ pre nejake $0 \leq t \leq s$. (Preco?) Lenze $t = 0$ by znamenalo $a^{*d} = 1$ v Z_n^*, co sme uz v nasom pripade vylucili. Teda $1 \leq t \leq s$. Polozme teraz $r = t - 1$. Zoberme

$$b = a^{*2^r d} \in Z_n^*.$$

Lahko sa ukaze, ze $b^{*2} = 1$ v Z_n^*, co je ekvivaletne s $b^2 \equiv 1 (\text{mod } n)$ v Z. Ktore prvky $l \in Z_n^*$ maju vlastnost $l^{*2} = 1$? Zrejme, $l^{*2} = 1$ je ekvivalentne s

$$l^2 - 1 \equiv 0 (\text{mod } n)$$

v Z. Posledny vyrok je ekvivalentny s

$$n \mid l^2 - 1 = (l - 1)(l + 1)$$

v Z. Pretoze n je prvocislo, tak dostaneme

$$n \mid l-1 \text{ alebo } n \mid l+1.$$

Pretoze $1 \leq l \leq n-1$, tak bud $l = 1$ alebo $l = n-1$. Ovsem $l = 1$ znamena $b = 1$, co neprichadza do uvahy, pretoze r bol najmensi exponent s vlastnostou $a^{*2^{(r+1)}d} = 1$. Ostava $b = n-1$. Zrejme,

$$b = a^{*2^r d} = n-1 \equiv -1 (\text{mod } n).$$

Priklad. Vezmime cislo $n = 561$. Chceme zistit, ci toto cislo je prvocislom, alebo nie. Podla vety 4 je 561 Carmichaelovym cislom, t.j. nevieme odpovedat na nasu otazku. Teraz pouzijeme vetu 5, teda Millerov-Rabinov test. Vyberme $a = 2$. Zrejme $\gcd(a, n) = 1$. Pomocou kalkulacky by sme zistili, ze $s = 4$, $d = 35$ a $2^{35} \equiv 263 (\text{mod } 561)$. Dalej vidime, ze $2^{2 \cdot 35} \equiv 166 (\text{mod } 561)$, $2^{4 \cdot 35} \equiv 67 (\text{mod } 561)$ a

$2^{8 \cdot 35} \equiv 1 (\mathrm{mod}\ 561)$. Teda, 561 nie je prvocislom, lebo nie je splnena ani druha podmienka vety 5.

Public-Key kryptosystemy (10. lekcia)

Podla doterajsich nasich skusenosti je pri tajnej komunikacii dvoch partnerov dolezite doruce-nie a utajenie kluca. Situacia sa este skomplikuje, ak mame n ucastnikov a kazdy z nich chce takto komunikovat s lubovolnym inym partnerom. V tomto pripade by sme museli tajne vymenit $n(n-1)/2$ klucov. Pri $n = 1000$ by to uz znamenalo vymenu 499500 klucov. Nepomoze ani vymena cez nejaku centralu. Potom by zase centrala mohla tajne sledovat komunikaciu. Preto vsetci ucastnici by mali mat doveru k centrale a ta by mala vediet dobre uschovat kluce. To tiez nie je prakticky uskutocnitelne. Odpoved je v asymetrickom kryptosysteme.

Diffieho-Hellmanov kryptosystem

Cvicenia. 1. Nech G je komutativna grupa a nech H a K su podgrupy grupy G. Dokazte, ze $H \cdot K = \{h \cdot k \in G : h \in H \text{ a } k \in K\}$ je podgrupa grupy G.

2. Nech H a K su podgrupy komutativnej grupy G. Utvorme priamy sucin grup $H \times K$. Dokazte, ze zobrazenie $\varphi : H \times K \to H \cdot K$ definovane predpisom $\varphi : (h, k) \mapsto h \cdot k$ je epimorfizmom. Dalej, ak $H \times K$ je cyklicka grupa, tak aj $H \cdot K$ je cyklickou grupou.

3. Nech $[g]$ je konecna cyklicka grupa a nech $\text{rad}(g) = n$. Predpokladajme, ze $k \mid n$. Dokazte, ze $\text{rad}(g^k) = n/k$.

4. Nech F je konecne pole. Dokazte, ze F obsahuje ako podpole Z_p pre nejake prvocislo

p. Zrejme, $p = \text{char}(F)$. (Z_p volame *prvotnym* podpolom pola F.)

Lema 1. Nech G je konecna komutativna grupa, pricom $g, h \in G$. Potom $[g] \times [h]$ je cyklicka grupa prave vtedy, ked

$$\gcd(\text{rad}(g), \text{rad}(h)) = 1.$$

Lema 2. Nech G je konecna komutativna grupa a nech $a, b \in G$. Potom $[a] \cdot [b]$ je cyklickou podgrupou grupy G, ked

$$\gcd(\text{rad}(a), \text{rad}(b)) = 1.$$

Dokaz. Predpokladajme, ze rady prvkov a a b su nesudelitelne. Potom vieme z lemy 1, ze $[a] \times [b]$ je cyklicka grupa. Podla cvic. 2 existuje epimorfizmus $\varphi : [a] \times [b] \to [a] \cdot [b]$, ktory zaruci, ze $[a] \cdot [b]$ je cyklickou grupou.

Veta 6. Nech $(F; +, \cdot)$ je konecne pole. Potom $F^* = (F - \{0\}; \cdot)$ je cyklicka grupa.

Dokaz. Cvic. 4 nam hovori, ze F je konecnorozmernym vektorovym priestorom nad svojim prvotnym podpolom Z_p. (Preco?) Ked to plati, tak mame v F konecnu bazu, povedzme $b_1, \cdots, b_n$ a preto lubovolny prvok $c \in F$ sa da jednoznacne napisat v tvare (linearna algebra!)

$$c = c_1 b_1 + \cdots + c_n b_n.$$

Z toho vidime, ze $\mid F \mid = p^n = q$. (Je to naozaj tak?) Dalej, F^* je komutativna grupa s $q - 1$ prvkami. Nech $F = \{0, a_1, \cdots, a_{q-1}\}$. Predpokladajme opak nasho tvrdenia, t.j. ze F^* **nie je cyklicka** grupa. Utvorme postupne vsetky cyklicke podgrupy $[a_1], \cdots, [a_{q-1}] \subseteq F^*$. Potom $[a_i] \neq F^*$, pre kazde $i = 1, \cdots, q - 1$. Vyberme podgrupu $[a_j]$ s **najvacsim** poctom prvkov k. Oznacenie: $a = a_j$. Podla nasho predpokladu,

$$\text{rad}(a) = k = \mid [a] \mid < q - 1.$$

Vezmime dalej lubovolny prvok $b \in F^*$. **Tvrdime,** ze rad$(b) \mid k$.

Dokazeme to nepriamo. Nech rad$(b) = l$ a nech l nedeli k. Cisla k a l sa daju prepisat do zovseobecneneho kanonickeho tvaru

$$k = p_1^{\alpha_1} \cdots p_r^{\alpha_r} \text{ a } \ l = p_1^{\beta_1} \cdots p_r^{\beta_r}.$$

(Pozor, $\alpha_i \geq 0$ a $\beta_i \geq 0$!) Potom, $l \mid k$ prave vtedy, ked $\beta_i \leq \alpha_i$ pre kazde $1 \leq i \leq r$. Podla posledneho predpokladu, existuje aspon jedno $1 \leq j \leq r$ take, ze

$$\beta_j > \alpha_j.$$

Kvoli jednoduchosti mozeme predpokladat, ze $j = 1$. Teda mame $\beta_1 > \alpha_1 \geq 0$. Ukazeme, ze to povedie k sporu. Vezmeme a^s, kde $s = p_1^{\alpha_1}$. Dalej b^m, kde $m = p_2^{\beta_2} \cdots p_r^{\beta_r}$. Zrejme, rad$(a^s) = k/s$ a rad$(b^m) = p_1^{\beta_1}$. Teda

$$\gcd(\text{rad}(a^s), \text{rad}(b^m)) = 1.$$

Ziskali sme novu cyklicku podgrupu $[a^s] \cdot [b^m]$ grupy F^* (lema 2). Jej rad sa rovna cislu $p_1^{(\beta_1 - \alpha_1)} \cdot k$, co je viacej ako k. Dostali sme spor s vyberom cisla k. Teda, $\beta_i \leq \alpha_i$ pre kazde $1 \leq i \leq r$. Inac povedane, ukazali sme, ze $\text{rad}(b) \mid k$ pre kazdy $b \in F^*$.

V poslednej casti dokazu vidime, ze lubovolny prvok $c \in F^*$ vyhovuje rovnici $x^k - 1 = 0$. S tym je ekvivaletny fakt, ze kazdy prvok $c \in F$ je riesenim rovnice $x^{k+1} - x = 0$. Dava to q rieseni. Na druhej strane, posledna rovnica moze mat nanajvys $k+1$ rieseni v poli F, co je spor, lebo $k + 1 < q$.

Diskretne logaritmy (11. lekcia)

V nasledujucom kryptosysteme z r. 1976 (Diffie, Hellman a spomina sa tiez Merkle) pojde o tajnu vymenu kluca pomocou verejne pristupnej siete. Cela vec spociva na komplikovanosti

matematickeho vypoctu tzv. diskretnych logaritmov.

Definicia. Nech $G = [g] = \{g^0, g^1, \cdots, g^{n-1}\}$ je konecna cyklicka grupa radu n. Nech $A \in G$. Potom existuje cislo $a \in \{0, 1, \cdots, n-1\}$ (jednoznacne urcene) s vlastnostou

$$A = g^a.$$

Cislo (mocninu) a volame *diskretnym logaritmom* prvku A pri zaklade g. Zapis: $a = \log_g(A)$.

Poznamka. 1. V kryptografii pouzivame casto cyklicku grupu $Z_p^* = (Z_p - \{0\}; \cdot)$ pre nejake prvocislo p. (Vid vetu 6.) 2. Uz zistovanie, ci $g \in Z_p^*$ je generatorom cyklickej grupy, alebo nie, moze byt komplikovanou zalezitostou.

Priklad. Nech $p = 13$. Da sa overit, ze Z_{13}^* ma styri generatory: 2, 6, 7 a 11. Podobne pre $p = 23$. Generatormi Z_{23}^* su 5, 7, 11 a 17.

Vymena kluca

Mame dvoch ucastnikov, ktori chcu tajne komunikovat. Najprv si cez verejnu siet (telefon, e-mail a pod.) vymenia udaje tykajuce sa kluca ich kryptosystemu. Dohodnu sa na nejakom prvocisle p a na nejakom generatori $g \in Z_p^*$. Zrejme, $2 \leq g \leq p-2$. (Preco?) To vsetko sa moze aj prezradit. Prvy ucastnik zvoli (uplne nahodne) cislo $a \in \{1, 2, \cdots, p-2\}$. Vypocita cislo

$$A \equiv g^a (\text{mod } p).$$

Cislo A posle cez verejnu siet druhemu ucastnikovi. Druhy ucastnik zvoli uplne nahodne cislo $b \in \{1, 2, \cdots, p-2\}$. Vypocita cislo

$$B \equiv g^b (\text{mod } p)$$

a posle cislo B prvemu ucastnikovi. Prvy ucastnik vypocita

$$B^a \equiv g^{ab} (\text{mod } p).$$

Zaroven druhy ucastnik pocita

$$A^b \equiv g^{ab} (\text{mod } p).$$

Teda klucom bude $K \in Z_p$, kde $K \equiv g^{ab} (\text{mod } p)$. Ku zvysku kryptosystemu sa este dostaneme.

Riziko ?

Pripadny spion sa moze dozvediet cisla p, g, A a B. Co nepozna, su $a = \log_g(A)$ a $b = \log_g(B)$. Potrebuje nejako zistit (vypocitat) kluc

$$K \equiv g^{ab} (\text{mod } p).$$

To sa zvykne volat *Diffieho-Hellmannov problem*. Zatial, jedine riesenie sa skyta cez prirodzene logaritmy. (Samozrejme az pri zvlastnych prvocislach, ked su dostatocne velke.) Pokial ich vypocet zostane komplikovany, bude obtiazne vypocitat aj kluc K. Inu metodu na vypocet K zatial nepozname.

Slabou strankou systemu sa moze stat skutocnost, ze nejaka tretia osoba, povedzme spion, sa bude vydavat za jedneho z ucastnikov. Nadviaze (diskretne) spojenie s prvym ako aj s druhym ucastnikom bez toho, zeby oni nieco vytusili. Dosiahne toho, ze obaja ucastnici budu komunikovat cez spiona. Presnejsie, spion si vymeni tajny kluc tak s prvym ako aj s druhym ucastnikom. Nasi ucastnici netusia, ze komunikuju cez spiona. Spion totiz prijme zakodovanu spravu od prveho ucastnika. Desifruje ju a zasifruje ju podla druheho kluca pre druheho ucastnika. Podobne to ide aj obratene, od druheho ucastnika ku prvemu.

Teraz sa pozrieme na odpovedajuci kryptosystem. Najprv zvolime vhodne prvocislo p. (Prakticky sa hladaju prvocisla, ktore maju aspon 512 bitov.) Potom polozime $P{=}C{=}Z_p$. Nech $M = m_1 m_2 \cdots m_s$ je binarna sprava, pricom m_i je binarne podslovo (blok) maximalne vzhladom na vlastnost, ze $m_i \in Z_p$ pre $1 \leq i \leq s$.

Najprv opiseme **sifrovanie**. Prvy ucastnik v ramci urcovania kluca urcil trojicu (p, g, A) a dozvedel sa od druheho ucastnika o cisle B. To su vsetko prvky Z_p. Tento prvy ucastnik postupne vypocita prvok

$$c_i = A^b * m_i = K * m_i \in Z_p,$$

pre kazde $1 \leq i \leq s$, kde K je kluc. Dostaneme zasifrovanu spravu $c = c_1 c_2 \cdots c_s$, ktora je napisana nad C, t.j. nad Z_p. Teda, ku klucu K, ktory je urceny stvoricou (p, g, A, B), ziskame sifrovaciu funkciu

$$E_K : M = m_1 \cdots m_s \mapsto c = c_1 \cdots c_s.$$

Pozrieme sa na to, ako tuto spravu mozeme **desifrovat**. Pretoze Z_p je pole a $0 \neq K \in Z_p$, tak existuje v Z_p prvok K^{-1} a mozeme napisat

$$m_i = K^{-1} * c_i,$$

zase pre vsetky $1 \leq i \leq s$. Takto sme dostali desifrovaciu funkciu

$$D_K : c = c_1 \cdots c_s \mapsto M = m_1 \cdots m_s.$$

RSA kryptosystem (12. lekcia)

Tento kryptosystem je pomenovany po Ron Rivestovi, Adi Shamirovi a Len Adlemanovi a publikovali ho r. 1978. Este aj dnes je to najdolezitejsi "public-key" kryptosystem. Jedna sa o asymetricky system. Kluc pozostava z dvoch casti: z verejnej (co sa moze zverejnit) a zo sukromnej (private), co zostava tajnym. Metoda je zalozena na teorii cisel a hlavne na jej dvoch rysoch: a) existuje rychly algoritmus na hladanie nahodnych prvocisel a b) neexistuje (presnejsie: nie je doposial znamy) rychly algoritmus na kanonicky rozklad velkeho cisla.

Zacneme vyrobou **kluca**. Povedzme, ze chcem prijimat zasifrovane spravy. Tak vyberiem dve

rozne prvocisla p a q. (Z praktickych dovodov ziadame, aby ich binarne rozklady boli aspon 512 bitove.) **Utvorim** cislo $n = p \cdot q$. Dalej **zvolim** cislo e, ktore vyhovuje vztahom

$$1 < e < \varphi(n) \text{ a } \quad \text{gcd(e, (p} - 1)(\text{q} - 1)) = 1,$$

kde φ je znama Eulerova funkcia a $\varphi(m)$ je pocet prirodzenych cisel $t \leq m$ s vlastnostou $\gcd(t, m) = 1$. Je zname, ze

$$\varphi(n) = \varphi(p) \cdot \varphi(q) = (p-1)(q-1).$$

Nakoniec, **vypocitam** cislo d, ktore splna

$$1 < d < \varphi(n) \text{ a } \quad de \equiv 1(\text{mod}(\varphi(n))$$

(veta 1). Zrejme e je neparne cislo.

Odkial vieme, ze take cisla e a d existuju? Vidime to z nasledujucej uvahy: Mozeme predpokladat, ze $3 < p < q$. Tvrdime, ze napr. $e = q$ je mozne. Naozaj,

$$q = (q-1)+1 < 2(q-1) < (p-1)(q-1) = \varphi(n)$$

a sucasne

$$\gcd(q, (p-1)(q-1)) = 1.$$

Vzhladom na posledny vztah vidime z vety 1, ze existuje $d \in Z^*_{\varphi(n)}$ s vlastnostou:

$$d * q = 1$$

v $Z^*_{\varphi(n)}$, co je pozadovane cislo.

Potom dvojica (n, e) predstavuje *verejny kluc*, t.j. moze sa zverejnit a d zostane utajene (u mna) a je to *sukromny kluc*. (Mame asymetricnost systemu.)
Priklad. Za prvocisla zvolme $p = 11$ a $q = 23$. Potom $n = 253$ a $(p-1)(q-1) = 4 \cdot 5 \cdot 11 = 220$. Cislo e mozeme zvolit male, t.j. $e = 3$. Z toho vypocitame zase $d = 147$. (Nie je to typicky priklad. Vzhladom na to, ze pouzite cisla su male, nevidno dovod, preco by sa nedal vypocitat sukromny kluc z verejneho.)

Dalsie casti kryptosystemu su P=C=Z_n. Potrebujeme este urcit sifrovaciu a desifrovaciu funkciu.

Sifrovacia funkcia

Predpokladajme, ze $M = m_1 \cdot m_2 \cdots m_s$ je binarna sprava, pricom m_i je binarne podslovo (blok) maximalne vzhladom na vlastnost, ze $m_i \in Z_n$, pre vsetky $1 \leq i \leq s$. (Tato poziadavka sa da aj nieco pozmenit: vsetky pouzite podslova m_i maju rovnaku dlzku a su z okruhu Z_n.) K binarnemu slovu M vyrobime nove binarne slovo $c = c_1 \cdot \cdots \cdot c_s$ nasledovne:

$$c_i = m_i^e (\text{mod } n)$$

pre kazde $1 \leq i \leq n$. Potom

$$E_{(n,e)} : M = m_1 \cdots m_s \mapsto c = c_1 \cdots c_s$$

je sifrovacia funkcia.

Desifrovacia funkcia

Lema 1. Nech $(Z_n; \oplus, *)$ je okruh zvyskovych tried modulo n. Nech

$$Z_n^* = \{a \in Z_n : \text{gcd}(a, n) = 1 \text{ v } Z\}.$$

Potom $(Z_n^*; *)$ je podgrupa pologrupy $(Z_n; *)$ a plati: $\mid Z_n^* \mid = \varphi(n)$.

Dokaz vyplyva z vety 1.

Lema 2. Nech $(Z_n^*; *)$ je grupa z lemy 1. Nech $a \in Z_n^*$. Potom

$$a^{*\varphi(n)} = 1$$

v Z_n^*.

Dokaz vyplyva z vety 1 a lemy 1.

Veta 7. Nech (n, e) je verejny a d sukromny kluc nejakeho RSA-kryptosystemu. Potom plati

$$(m^e)^d \equiv m (\text{mod } n)$$

pre kazde prirodzene cislo m. Inac povedane, v $(Z_n; *)$ plati

$$m^{*ed} = m$$

pre kazde $m \in Z_n$.

Dokaz. Podla predpokladu je

$$ed \equiv 1 (\text{mod } \varphi(n)).$$

To je ekvivalentne s vyrokom, ze

$$ed = t\varphi(n) + 1$$

pre vhodne $t \in Z$, pricom vieme, ze $\varphi(n) = (p-1)(q-1)$. Potom

$$(m^e)^d = m^{ed} = m^{1+t\varphi(n)} = m(m^{\varphi(n)})^t$$

v Z. Z tychto rovnosti dostaneme

$$(m^e)^d = m^{ed} = m(m^{p-1})^{(q-1)t} \equiv m (\text{mod } p)$$

na zaklade Malej Fermatovej vety. Podobne

$$(m^e)^d = m^{ed} \equiv m (\text{mod } q).$$

Posledne dva vztahy su logicky ekvivalentne s vyrokmi: $p \mid m^{ed} - m$ a $q \mid m^{ed} - m$ v Z. Pretoze $p \neq q$ su prvocisla, tak $\gcd(p, q) = 1$. Z toho vyplyva (podla znameho tvrdenia), ze

$$n = p \cdot q \mid m^{ed} - m.$$

Posledny vztah znamena

$$(m^e)^d \equiv m (\text{mod } n)$$

pre kazde $m \in Z$, z coho uz dostaneme nase tvrdenie.

Teraz vidime, ze

$$D_d : c = c_1 \cdots c_s \mapsto D_d(c_1) \cdots D_d(c_s) = c_1^d \cdots c_s^d$$

je desifrovacia funkcia, pretoze

$$D_d(c_i) = c_i^d = (m_i^e)^d = m_i$$

pre kazde $1 \leq i \leq s$ (podla vety 7).