UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
Podiplomski študij matematike - izobraževalna smer

Tečaj iz kriptografije in teorije kodiranja - 2003/04
Predavatelj: Aleksandar Jurišić

IDENTIFIKACIJSKA ŠTEVILA

(seminarska naloga)

TATJANA TOMAŠ

Kazalo

Povzetek

Identifikacijska števila so povsod okoli nas. V seminarski nalogi so le-ta najprej predstavljena. Pri prenosu identifikacijskih števil pride pogosto do napak. Da bi se te napake odkrile, se identifikacijskim številom dodajajo kontrolne številke. Predstavljena so različna identifikacijska števila in načini izračuna kontrolne številke. Na koncu pa je še na primeru opisana zgradba črtne kode.

1 Identifikacijska števila in kontrolne številke

1.1 Predstavitev identifikacijskih števil

Kamor koli se ozremo, so okoli nas same številke, ki se jih ponavadi sploh ne zavedamo. Gremo v trgovino in vsak izdelek ima svojo številko - tisto, ki jo prodajalka na blagajni prek črtne kode vnese v računalnik. Plačamo s plačilno kartico, na kateri je številka našega transakcijskega računa. Ko gremo v zdravstveni dom, potrebujemo zdravstveno kartico, na kateri je številka našega zdravstvenega zavarovanja. Za našo identifikacijo potrebujemo osebno zkaznico in na nekaterih mejah potni list. Na obeh imamo OCR-B zapis, v katerem so naši podatki, ki se berejo s posebnimi čitalci. Seveda oba dokumenta vsebujeta naš EMŠO (enotna matična številka občana). Večina teh številk predstavlja neko informacijo in vsa ta števila so primeri identifikacijskih števil.

Identifikacijska števila so primeri kod - skupina simbolov, ki predstavlja informacijo. Uporabljajo se torej za identifikacijo raznih stvari, na primer izdelkov v trgovini, ljudi, dokumentov, računov. V današnji dobi informacij z visoko zmogljivimi računalniki so uporabljena za hranjenje in lažje razvrščanje informacij, ki so ponavadi zakodirane s črtnimi kodami, oziroma za lažjo identifikacijo izdelkov, dokumentov, računov...

1.2 Tipi identifikacijskih števil

Identifikacijska števila so ponavadi v obliki zaporedja števil, črk, simbolov ali kombinacije vseh treh. Glede na njihov namen, se izbira tudi njihova oblika. Najpogostejše so tri oblike:

- *identifikacijska števila sestavljena iz samih števk (0, 1, ..., 9)*:
  npr.: EMŠO 1804978505088, EAN-13 3838800000756,

- *identifikacijska števila sestavljena iz števk in črk (lahko ločene s črtico)*:
  npr.: številka potnega lista P00563284, ISBN 0-201-52032-X,

- *identifikacijska števila sestavljena iz števk, črk in simbolov (*, /, #, ...)*:
  npr.: številka vozniškega dovoljenja (ZDA) MOE**TH220DW.

1.3 Napake pri prenosu identifikacijskih števil

Identifikacijska števila se vsakodnevno "prenašajo" - posredujemo jih ustno naposredno ali preko telefona, pisno, preko interneta, vnašamo jih v računalnik ... Pri vsakem prenosu se lahko zgodi, da pride do napake, da na primer zamenjamo dve števki. Tako lahko pride v nekaterih primerih do hudih napak, na primer pri številkah transakcijskih računov, kjer bi bil denar nakazan na napačen račun, ali pa na blagajni v trgovini, kjer se cene na izdelkih in odčitane na blagajni ne bi ujemale.

Najbolj tipične napake pri prenosu identifikacijskih števil so naslednje [2]:

- *napaka ene števke* $(a \to b)$: ena števka v številu se spremeni,
- *zamenjava sosednjih števk* $(ab \to ba)$: števki, ki stojita skupaj, se zamenjata,
- *"skok" zamenjava* $(abc \to cba)$: števki, ki sta ločeni z eno števko, se zamenjata,
- *dvojna zamenjava* $(aa \to bb)$: par enakih števk se spremeni v drug par enakih števk,
- *fonetična zamenjava* $(a0 \leftrightarrow 1a,\ a = 2, 3, \ldots, 9)$: ta napaka je značilna za angleški jezik, na primer 14 (fourteen) se sliši kot 40 (forty),
- *"skok" zamenjava dveh enakih števk* $(aca \to bcb)$: dve enaki števki, ločeni z eno števko, se spremenita v dve drugi enaki števki.

Tabela 2. 1 prikazuje primere napak in pogostost napak glede na vse napake pri prenašanju [2].

tip napake	število	napaka	relativna frekvenca
napaka ene števke	191433	191933	79,1 %
zamenjava sosednjih števk	191433	191343	10,2 %
"skok" zamenjava	191433	193413	0,8 %
dvojna zamenjava	191433	191455	0,5 %
fonetična zamenjava	191433	194033	0,5 %
"skok" zamenjava dveh enakih števk	191433	393433	0,3 %

Tabela 2. 1: Primeri napak in pogostost napak pri prenosu identifikacijskih števil.

1.4 Dodajanje kontrolne številke

Ker prihaja do napak pri prenosu identifikacijskih števil, torej do napak pri nepravilnem vnašanju številk v računalnik ali pri nepravilnem odčitavanju črtnih kod ali drugih zapisov (OCR-B) ali pri drugačnih prenašanjih identifikacijskih števil, je treba poskrbeti, da so identifikacijska števila pravilno prenešena oziroma, da če pride do napake, da se ta tudi odkrije. Če je možno pa tudi, da se ta napaka odpravi avtomatično.

To privede do razvoja metod, s katerimi prejemnik prepozna, če je bila identifikacijska številka narobe prenesena oziroma, če je bila črtna koda (ali kak drug zapis, npr. OCR-B) narobe odčitana. Identifikacijskim številkam dodamo kontrolno številko. Metodo dodajanja kontrolne številke oziroma način izračuna kontrolne številke imenujemo **shema kontrolne številke**.

Cilj določanja kontrolne številke je torej ujeti vse prej omenjene tipične napake oziroma vsaj prvi dve. Najpogosteje pri izračunu kontrolne številke uporabimo teorijo šteil, redkeje pa permutacije in teorijo grup.

Naj bo **identifikacijska shema** skupno ime za način določanja identifikacijskega števila in način izračuna kontrolne številke (shema kontrolne številke). Treba je še poudariti, da ko določene informacije preoblikujemo v identifikacijsko število, da dve različni informaciji ne smeta imeti enako identifikacijsko številko. Torej mora biti preslikava [1], ki slika iz določenih informacij v identifikacijska števila injektivna.

V naslednjih štirih poglavjih si bomo ogledali nekaj identifikacijskih shem (iz [2], [4] in primeri pri nas) in ugotavljali, kako so te sheme učinkovite oziroma katere od teh zaznajo napako ene številke in zamenjavo sosednjih številk.

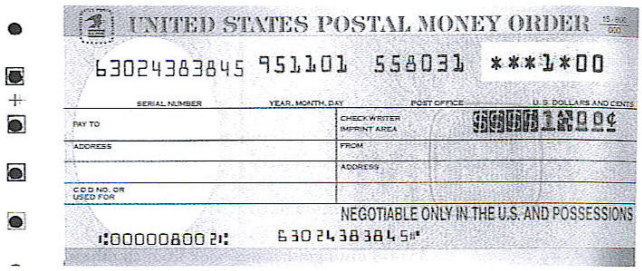
[1] zgoščevalna funkcija - poljubno dolgemu sporočilu priredi kratko zaporedje bitov; zgoščevalna funkcija brez trčenj - za $x \neq x'$ velja $H(x) \neq H(x')$

2 Identifikacijske sheme in teorija števil

2.1 US Postal Money Orders

Začnimo s poštno denarno nakaznico, ki jo uporablja US Post Office, torej pošta v ZDA. Identifikacijska številka je enajstmestna $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}$. Prvih deset števk je številka nakaznice, enajsta pa kontrolna številka, ki je enaka ostanku pri deljenju vsote prvih deset števk z 9:

$$a_{11} = (a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + a_9 + a_{10}) \pmod 9$$



Slika 3. 1: Poštna denarna nakaznica v ZDA (številka spodaj je "berljiva" z računalnikom).

Identifikacijska številka iz slike 3. 1 je 63024383845 :

$$a_{11} = (6+3+0+2+4+3+8+3+8+4) \pmod 9 = 41 \pmod 9 = 5$$

Ker je vsota komutativna, ta shema ne prepozna zamenjave sosednjih števk, razen, če je v tej zamenjavi kontrolna številka. Prepozna pa skoraj vse napake ene števke, razen če v prvih desetih števkah namesto 0 zapišemo 9 in obratno, kar sledi iz lastnosti pri računanju z moduli.

Pri potovalnih čekih American Express in VISA je prav tako uporabljeno deljenje z 9. V teh dveh primerih je kontrolna številka izbrana tako, da je vsota vseh števk, vključno s kontrolno številko, deljiva z 9.

Slika 3. 2: Potovalni ček z identifikacijsko številko 3875050557 (kontrolna številka je 7).

2.2 Letalska karta (ZDA)

Identifikacijska številka letalske karte v ZDA je petnajstmestna. Prva števka je številka kupona: številka 1 označuje prvi let, številka 2 drugi let, ... Številka kupona 0 pa označuje strankino potrdilo o prejemu. Druga, tretja in četrta števka so številka letala. Tretji del $a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}a_{14}$ pa je številka dokumenta. Zadnja števka je kontrolna številka, ki je enaka ostanku pri deljenju števila $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}a_{14}$ s 7:

$$a_{15} \;=\; a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}a_{14} \pmod 7$$

Slika 3. 3: Letalska karta v ZDA.

Identifikacijska številka iz slike 3. 3 je 0 012 7881879532 1:

$$a_{15} \;=\; 00127881879532 \pmod 7 \;=\; 127881879532 \pmod 7 \;=\; 1$$

Poglejmo si, katere napake prepozna ta shema.

Recimo, da v številu a' zamenjamo eno števko $a_i = a\,(\,i = 1, \ldots, 14\,)$ z b $(a \neq b)$. Imamo torej dve različni števili: $a' = a_1 \ldots a \ldots a_{15}$ in $b' = a_1 \ldots b \ldots a_{15}$, ki se razlikujeta v eni števki. Sistem ne bo zaznal napake, če

$$a' = b' \pmod 7 \quad \Leftrightarrow \quad |a' - b'| = 0 \pmod 7 \quad \Leftrightarrow$$

$$\Leftrightarrow \quad |(a-b)0 \ldots 0| = 0 \pmod 7 \quad \Leftrightarrow \quad |a-b| = 7\,.$$

Če spremenimo kontrolno številko, bo napaka seveda "ujeta", kar se vidi iz načina izračuna kontrolne številke.

Kaj pa zamenjava katerikolih dveh sosednjih števk ($\ldots ab \ldots \;\to\; \ldots ba \ldots$) razen, če v zamenjavi ni kontrolne številke? Za $a \neq b$ sistem ne bo zaznal zamenjave za števili $a' = a_1 \ldots ab \ldots a_{15}$ in $b' = a_1 \ldots ba \ldots a_{15}$, če

$$a' = b' \pmod 7 \quad \Leftrightarrow \quad |a' - b'| = 0 \pmod 7 \quad \Leftrightarrow$$

$$\Leftrightarrow \quad |(ab-ba)0 \ldots 0| = 0 \pmod 7 \quad \Leftrightarrow \quad |ab-ba| = 0 \pmod 7 \quad \Leftrightarrow$$

$$\Leftrightarrow \quad ab = ba \pmod 7 \quad \Leftrightarrow \quad ab \in \{07, 70, 18, 81, 29, 92\} \quad \Leftrightarrow \quad |a-b| = 7\,.$$

Če pa zamenjamo zadnji dve števki ($\ldots a_{14}a_{15} \;\to\; \ldots a_{15}a_{14}$), bo zamenjava zaznana.

2.3 Sistem EAN.UCC

Kot smo že omenili, imajo vsi artikli v trgovini svojo številko in črtno kodo. Osnove za te številke so standardi EAN.UCC [2]. Sistem EAN.UCC je zbir standardov, ki omogočajo učinkovito upravljanje preskrbovalne verige z enoličnim označevanjem proizvodov, transportnih enot, lokacij in storitev. Osnovo sistema predstavlja globalno enolična identifikacijska številka (GTIN [3]). Številko ali intervale številk GTIN lastniku blagovne znamke podeli nacionalno EAN združenje. Zaradi avtomatskega zajema podatkov se identifikacijska številka zapiše v obliki črtne kode. Identifikacijska številka je negovoreča [4], zato so vsi potrebni podatki zapisani v podatkovni bazi. GTIN je neodvisna od cen in načinov dobave in se pojavlja v katalogih, prospektih, cenikih in na dokumentih ali elektronskih poslovnih sporočilih.

V sistemu EAN.UCC so trije glavni elementi sistema oštevilčevanja:

- *globalno trgovinska identifikacijska številka (GTIN)*, ki se uporablja za edinstveno identifikacijo prodajnih enot [5] po vsem svetu. Identifikacija in označevanje prodajnih enot omogoča avtomatizacijo maloprodajnega mesta (s pomočjo datotečnih cenikov), prevzema proizvodov, upravljanja zalog, avtomatskega ponovnega naročanja, prodajne analize ter širokega področja drugih poslovnih aplikacij.
- *zaporedna koda zabojnika (SSCC)*, je številka oziroma podatkovna struktura, ki se uporablja za edinstveno (enoznačno) identifikacijo logističnih enot [6].
- *globalna lokacijska številka (GLN)*, ki se uporablja za identifikacijo podjetja ali organizacije.

Prodajne enote se oštevilčujejo z GTIN s pomočjo štirih sistemov oštevilčevanja: **EAN.UCC - 8**, **UCC - 12**, **EAN.UCC - 13** in **EAN.UCC - 14**. Če so vključeni v podatkovno bazo, so vsi shranjeni v 14-mestnem polju. Izbira sistema je odvisna od narave enote in od obsega uporabniških aplikacij. Identifikacijske številke so zapisane tudi v črtnih kodah. V sistemu EAN.UCC se uporabljajo tri različne simbologije črtne kode:

- EAN/UPC [7]: EAN - 13, UPC - A ali EAN - 8, UPC - E,
- ITF - 14,
- UCC.EAN - 128 [8].

[2] EAN - European Article Number, UCC - Uniform Code Concil

[3] GTIN - Global Trade Item Number

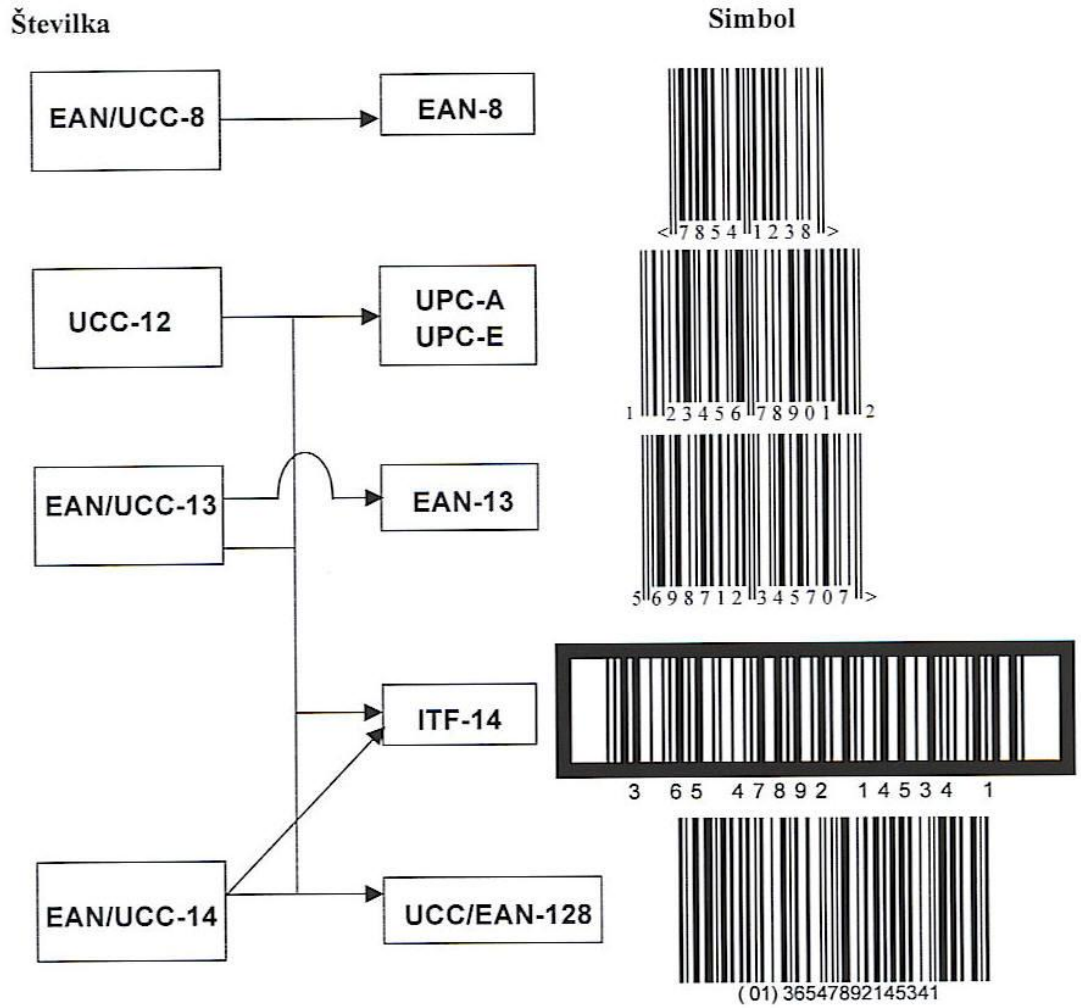
[4] v sami številki ni vsebovana nobena posebna informacija

[5] Prodajna enota je katerakoli postavka (proizvod ali storitev), o kateri je treba poiskati kako vnaprej določeno informacijo in ki se lahko ovrednoti, naroči ali fakturira v katerikoli točki katerekoli preskrbovalne verige.

[6] Logistična enota je enota poljubne sestave, pripravljena za transport in/ali skladiščenje, ki jo je treba upravljati (spremljati v preskrbovalni verigi).

[7] UPC - Universal Product Code

[8] Edina simbologija EAN.UCC, ki omogoča kodiranje dodatnih informacij poleg identifikacije je UCC.EAN - 128.



Slika 3. 4: Izbira črtne kode glede na identifikacijsko številko.

Glavna uporaba sistema EAN.UCC je identifikacija artiklov, ki so namenjeni za odčitavanje (skeniranje) na maloprodajnem mestu in se imenujejo tudi potrošniške enote. Slednje se identificirajo s številko EAN.UCC-13 (ali s številko UCC-12, če se prodajajo v Severni Ameriki), če so zelo majhne, pa s številko EAN.UCC-8 (oziroma z UCC-12 brez ničel).

Pa si poglejmo strukture oštevilčenja, s katerimi se srečujemo potrošniki: identifikacijske številke pod črtnimi kodami EAN-13, UPC-A ali EAN-8, UPC-E.[9]

[9]Pogledali si bomo, kako so sestavljene najpogostejše identifikacijske številke pod omenjenimi črtnimi kodami. Izjeme kot npr. prodajne enote s spremenljivo vsebino (sadje, zelenjava, meso, vrvi, tkanine,...) bomo izpustili. Poseben primer kot so knjige pa bomo pogledali v naslednjem podpoglavju.

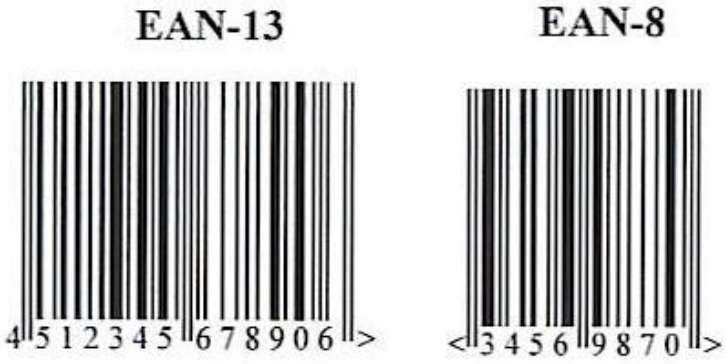
Najprej poglejmo sistem v Severni Ameriki[10] (ZDA in Kanada). Obstaja pet verzij UPC, najpogosteje pa se uporabljata dve: UPC - A in UPC - E.

Slika 3. 5: Črtni kodi UPC - A in UPC - E .

Številka pod črtno kodo UPC - A je dvanajstmestna $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}$. Prva števka a_1 označuje tip artikla (0 - špecerija, 2 - meso in mesni izdelki, 3 - zdravila in izdelki za zdravje, 4 - izdelki, ki niso hrana, 5 - karte, kuponi, 6, 7 - drugi artikli, 1, 8, 9 - rezervirano za nadaljnjo uporabo). Niz števk $a_2a_3a_4a_5a_6$ označuje proizvajalca. Števke $a_7a_8a_9a_{10}a_{11}$ označujejo artikel. Zadnja števka a_{12} pa je kontrolna številka[11]. Če ima proizvajalec veliko artiklov, je lahko njegova številka krajša, zato da več števil ostane za označitev artiklov (npr. $a_2a_3a_4a_5$ številka proizvajalca, $a_6a_7a_8a_9a_{10}a_{11}$ številka artikla).

Verzija UPC - E je osemmestna in se uporablja za majhne artikle, kjer ni mogoča uporaba UPC - A.

Poglejmo še sestavo identifikacijskih številk pod črtnima kodama EAN - 13 in EAN - 8.



Slika 3. 6: Črtni kodi EAN - 13 in EAN - 8.

[10]Mnogi severnoameriški uporabniki še ne morejo prilagoditi svojih datotek za delo z identifikacijskimi številkami EAN.UCC - 13. Čeprav je bil za globalni prevzem EAN.UCC - 13 določen datum v januarju 2005, standardna struktura oštevilčenja UCC - 12, predstavljena v simbolih UPC - A ali UPC - E je potrebna do tedaj. Podjetja, ki prodajajo svoje blago na ameriških in kanadskih trgih morajo od svoje članske organizacije UCC zaprositi za predpono podjetja, ki jo vodi UCC.

[11]Kontrolna številka ni vedno izpisana pod črtno kodo, je pa vedno zakodirana v črtni kodi.

Številka pod črtno kodo EAN - 13 je trinajstmestna $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}$. Prvi dve ali tri številke $a_1a_2a_3$ tvorijo predpono EAN.UCC, katero sourejata EAN International in UCC. To ne pomeni vedno, da je artikel proizveden ali distribuiran v državi kodiranja. Predpona EAN.UCC samo označuje državo nacionalne organizacije EAN [12], ki je podjetje včlanila in ji dodelila številko. Sledi številka podjetja EAN.UCC, ki jo določi nacionalna organizacija EAN ali UCC. Predpona EAN.UCC in številka podjetja tvorita predpono podjetja EAN.UCC. Na splošno vsebuje šest do deset številk, odvisno od potrebe podjetja. Naslednji niz številk, ponavadi od ene do šest številk, označuje artikel [13] (to število določi podjetje samo). Zadnja številka a_{13} pa je kontrolna številka. Skrajšano verzijo EAN - 13, to je EAN - 8, podeljuje pri nas EAN Slovenija za vsak artikel posebej.

Poglejmo si sedaj shemo kontrolne številke.

Standardni izračun kontrolne cifre za strukture oštevilčevanja EAN.UCC

	Pozicije cifer																	
EAN/ UCC-8											N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
UCC-12							N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}
EAN/ UCC-13						N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}
EAN/ UCC-14					N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}	N_{14}
	Pomnoži vrednost vsake pozicije z																	
	x3	x1	x3	x1	x3	x1	x3	x1	x3	x1	x3	x1	x3	x1	x3	x1	x3	
	Akumulirani rezultati = *vsota*																	
	Odštej *vsoto* od najbližjega večkratnika 10 = *kontrolna cifra* →																	

Slika 3. 7: Izračun kontrolne številke.

Za identifikacijsko številko pod UPC - A torej velja:

$$3 \cdot \sum_{i=2k-1,\, k\in\mathbb{N}}^{11} a_i + \sum_{i=2k,\, k\in\mathbb{N}}^{12} a_i = 0 \pmod{10}.$$

Če gledamo kot množenje vektorjev, lahko zgornji zapis zapišemo drugače:

$$(3,1,3,1,3,1,3,1,3,1,3,1)\cdot(a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11},a_{12}) = 0 \pmod{10}.$$

Kontrolno številko lahko torej izračunamo po formuli:

$$a_{12} = \left(10 - \left(3 \cdot \sum_{i=2k-1,\, k\in\mathbb{N}}^{11} a_i + \sum_{i=2k,\, k\in\mathbb{N}}^{10} a_i\right) \bmod 10\right) \bmod 10.$$

[12] Slovensko združenje EAN Slovenija ima kodo 383.

[13] Ta niz imenujemo referenca prodajne enote in je nepomensko število, kar pomeni, da se posamezne številke v tem številu ne nanašajo na nobeno klasifikacijo oziroma ne nosijo nobene določene informacije.

Identifikacijska številka UPC - A iz slike 3. 5 je 0 53600 10054 0 :

$$\begin{aligned} a_{12} &= (\,10 - (\,3 \cdot (\,0+3+0+1+0+4\,) + 5 + 6 + 0 + 0 + 5\,) \bmod 10) \bmod 10 = \\ &= (\,10 - 40 \bmod 10\,) \bmod 10 = \\ &= 10 \bmod 10 = \\ &= 0. \end{aligned}$$

Podobno velja za identifikacijsko številko pod EAN - 13:

$$\sum_{i=2k-1,\, k\in\mathbb{N}}^{13} a_i + 3 \cdot \sum_{i=2k,\, k\in\mathbb{N}}^{12} a_i = 0 \pmod{10}$$

oziroma

$$(1,3,1,3,1,3,1,3,1,3,1,3,1)\cdot(a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11},a_{12},a_{13}) = 0 \pmod{10}.$$

Kontrolno številko lahko torej izračunamo po formuli:

$$a_{13} = \left(10 - \left(\sum_{i=2k-1,\, k\in\mathbb{N}}^{11} a_i + 3 \cdot \sum_{i=2k,\, k\in\mathbb{N}}^{12} a_i\right) \bmod 10\right) \bmod 10.$$

Identifikacijska številka EAN - 13 iz slike 3. 6 je 4 512345 678906 :

$$\begin{aligned} a_{13} &= (\,10 - (\,4+1+3+5+7+9+3\cdot(\,5+2+4+6+8+0\,)\,) \bmod 10) \bmod 10 = \\ &= (\,10 - 104 \bmod 10\,) \bmod 10 = \\ &= 6 \bmod 10 = \\ &= 6. \end{aligned}$$

Poglejmo si idetifikacijsko shemo za EAN - 13. Katere napake prepozna ta shema? Ta shema je boljša od prejšnjih dveh ("mod7", "mod9"), saj prepozna vse napake ene števke.

Recimo, da v identifikacijski številki $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}$ zamenjamo a_i, $1 \le i \le 13$, z a_i' ($a_i \neq a_i'$). Označimo

$$a = (1,3,1,3,1,3,1,3,1,3,1,3,1)\cdot(a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11},a_{12},a_{13})\,.$$

Pri izračunu števila a' vzamemo izračun a, kjer samo spremenimo a_i v a_i'. Odštejmo $a - a'$. Napak bo zaznana samo v primeru $a - a' \neq 0 \pmod{10}$. Ločimo dva primera:

1. i je liho število:
$$a - a' = a_i - a_i' \neq 0 \pmod{10}\,,$$

2. i je sodo število:

$$a - a' = 3\,a_i - 3\,a_i' = 3\,(a_i - a_i') \neq 0 \pmod{10}\,.$$

Vidimo torej, da ta shema res prepozna vse napake ene števke.

Poglejmo še, kaj se zgodi, če v številu $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}$ zamenjamo sosednji števki, recimo a_i in a_{i+1}, $1 \leq i \leq 12$. Izračunamo a in a' (podobno kot zgoraj). Odštejmo $a - a'$. Napaka bo odkrita samo v primeru $a - a' \neq 0 \pmod{10}$. Ločimo dva primera:

1. i je liho število:

$$a - a' = a_i + 3\,a_{i+1} - a_{i+1} - 3\,a_i = 2\,a_{i+1} - 2\,a_i = 2\,(a_{i+1} - a_i)\,,$$

2. i je sodo število:

$$a - a' = 3\,a_i + a_{i+1} - 3\,a_{i+1} - a_i = 2\,a_i - 2\,a_{i+1} = 2\,(a_i - a_{i+1})\,.$$

Če želimo, da bo $a - a' \neq 0 \pmod{10}$, mora biti $|a_i - a_{i+1}| \neq 5$.

Ta shema torej odkrije zamenjavo sosednjih števk ($a_i\,a_{i+1} \rightarrow a_{i+1}\,a_i$), če je

$$|a_i - a_{i+1}| \neq 5\,.$$

2.4 ISBN

Na vseh novejših knjigah najdemo desetmestno številko ISBN [14]. Deset števk oznake ISBN je razdeljenih na 4 elemente spremenljivih dolžin, ki morajo biti med seboj ločeni z vezajem ali presledkom [15]. Npr.: ISBN 0 571 08989 5 ali 90-70002-34-5. Število znakov v prvih treh elementih oznake ISBN (oznaka skupine, oznaka založnika in oznaka naslova) se spreminja. Število znakov v oznaki skupine in oznaki založnika je določeno glede na število naslovov, kij jih založnik ali založniška skupina namerava izdati. Založniki in založniške skupine z velikim številom izdanih naslovov imajo krajše oznake.

Sestava ISBN:

- *oznaka skupine:* Prvi element oznake ISBN označuje državo, geografsko ali jezikovno področje, ki sodeluje v sistemu ISBN. Nekateri člani tvorijo jezikovno področje (npr. oznaka skupine 0 - angleško jezikovno področje, 2 - francosko jezikovno področje, 3 - nemško jezikovno področje, ...) ali geografsko področje (npr. Južni Pacifik - oznaka skupine 982). Oznaka skupine ima lahko do 5 znakov. Vse oznake skupin dodeljuje Mednarodna agencija za ISBN v Berlinu.

- *oznaka založnika:* Drugi element oznake ISBN označuje založnika. Oznaka založnika navadno natančno identificira založniško hišo in njen naslov. Ko založniki izkoristijo vse svoje proste oznake naslovov, lahko dobijo novo identifikacijsko oznako. Oznaka založnika ima lahko do 7 znakov. Oznake založnika dodeljujejo agencije, ki so zadolžene za upravljanje z oznakami ISBN znotraj države, geografskega ali jezikovnega področja, kjer imajo posamezni založniki tudi svoj sedež.

- *oznaka naslova:* Tretji element oznake ISBN označuje določeno izdajo publikacije pri določenem založniku. Oznaka naslova lahko ima do 6 znakov. Ker mora imeti celotna oznaka ISBN vedno 10 znakov, so nezasedena mesta dopolnjena z vodilnimi ničlami.

- *kontrolna številka:* Zadnji znak oznake ISBN je kontrolna številka. Izračunana je po modulu 11 z utežmi 10 - 2, pri čemer se uporabi X namesto 10, kadar je izračunana kontrolna številka 10. To pomeni, da vsako od prvih devetih števk oznake ISBN (seveda brez kontrolne številke) pomnožimo po vrsti s številko od 10 do 2. Vsota tako dobljenih produktov in kontrolne številke mora biti deljiva z 11 brez ostanka.

Shema kontrolne številke desetmestne $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}$ ISBN je torej naslednja:

$$10\,a_1 + 9\,a_2 + 8\,a_3 + 7\,a_4 + 6\,a_5 + 5\,a_6 + 4\,a_7 + 3\,a_8 + 2\,a_9 + a_{10} = 0 \pmod{11}$$

oziroma

$$(10, 9, 8, 7, 6, 5, 4, 3, 2, 1) \cdot (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}) = 0 \pmod{11}.$$

[14] ISBN - International Standard Book Number oz. mednarodna standardna knjižna številka

[15] Za računalniško obdelavo podatkov se uporablja niz desetih števk brez vezajev ali presledkov. Interpretacija in človeku namenjeni izpis temeljita na izpostavljeni oznaki skupine in oznaki založnika.

Kontrolno številko lahko torej izračunamo po formuli:

$$a_{10} = (11 - (10\,a_1 + 9\,a_2 + 8\,a_3 + 7\,a_4 + 6\,a_5 + 5\,a_6 + 4\,a_7 + 3\,a_8 + 2\,a_9) \bmod 11) \bmod 11.$$

Če dobimo $a_{10} = 10 \quad \Rightarrow \quad a_{10} \equiv X$.

Poglejmo si izračun kontrolne številke na dveh primerih:

1. ISBN 0-901690-54-6 :

$$10 \cdot 0 + 9 \cdot 9 + 8 \cdot 0 + 7 \cdot 1 + 6 \cdot 6 + 5 \cdot 9 + 4 \cdot 0 + 3 \cdot 5 + 2 \cdot 4 = 192 \quad \Rightarrow$$

$$\Rightarrow \quad a_{10} = (\,11 - 192 \bmod 11\,) \bmod 11 = 6 \bmod 11 = 6\,,$$

2. ISBN 0-201-52032-X :

$$11 - (\,10 \cdot 0 + 9 \cdot 2 + 8 \cdot 0 + 7 \cdot 1 + 6 \cdot 5 + 5 \cdot 2 + 4 \cdot 0 + 3 \cdot 3 + 2 \cdot 2 = 78 \quad \Rightarrow$$

$$\Rightarrow \quad a_{10} = (\,11 - 78 \bmod 11\,) \bmod 11 = 10 \bmod 11 = 10\,.$$

Prišli smo do prvega primera, kjer je identifikacijsko število (ISBN) sestavljeno tako, da se ugotovijo vse napake ene števke in vse zamenjave dveh sosednjih števk. Kljub tej prednosti pa ima koda ISBN vseeno dve slabosti: prva je ta, da lahko v kodi nastopi črka (X), druga pa ta, da je dolžina kode fiksna (desetmestno število). Kasneje bomo spoznali identifikacijske sheme, ki odkrijejo obe vrsti napak in nimajo omenjenih slabosti.

Pa pokažimo, zakaj ta shema zazna vse napake ene števke in vse zamenjave dveh sosednjih števk.

Recimo, da v identifikacijski številki $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}$ zamenjamo a_i, $1 \le i \le 10$, z a_i' ($a_i \neq a_i'$). Označimo

$$a = (10, 9, 8, 7, 6, 5, 4, 3, 2, 1) \cdot (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10})\,.$$

Pri izračunu števila a' vzamemo izračun a, kjer samo spremenimo a_i v a_i'. Odštejmo $a - a'$:

$$a - a' = k\,a_i - k\,a_i' = k \cdot (a_i - a_i'),\ k = 11 - i\,.$$

Ker sta k in 11 tuji števili in $a_i - a_i'$ ter 11 prav tako, velja $a - a' \neq 0 \pmod{11}$. Sledi torej, da je napaka odkrita.

Poglejmo še, kaj se zgodi, če v številu $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}$ zamenjamo sosednji števki, recimo a_i in a_{i+1}, $1 \le i \le 9$. Izračunamo a in a' (podobno kot zgoraj). Odštejmo $a - a'$:

$$\begin{aligned} a - a' &= k\,a_i + (k-1)\,a_{i+1} - k\,a_{i+1} - (k-1)\,a_i = \\ &= k\,a_i + k\,a_{i+1} - a_{i+1} - k\,a_{i+1} - k\,a_i + a_i = \\ &= a_i - a_{i+1} \neq \\ &\neq 0 \pmod{11}. \end{aligned}$$

Sledi torej, da je napaka odkrita.

Hitro in vsesplošno razširjeno računalniško branje črtnih kod je privedlo do dogovora med Mednarodnim združenjem za številčenje proizvodov (EAN), Uniform Code Council (UCC) in Mednarodno agencijo za ISBN, ki omogoča pretvorbo oznake ISBN v črtno kodo EAN. To seveda še povečuje moč sistema ISBN kot mednarodnega sistema identifikacije znotraj vsesplošne sheme črtnih kod.

Vse črtne kode EAN se začenjajo z oznako za nacionalno identifikacijo, razen kod za označevanje publikacij. Prej omenjeni sporazum nadomešča oznako nacionalne identifikacije s posebno oznako "območje knjig" (Bookland), ki ga predstavljajo številke 978 za knjige. EAN oznaki "območje knjig" 978 sledi prvih devet številk oznake ISBN. Kontrolno številko v ISBN izpustimo in jo nadomestimo z EAN kontrolno številko, izračunano po pravilih EAN (modul 10). Na razpolago je tudi dodatna petmestna koda, ki jo lahko uporabimo za druge dodatne informacije, npr. o priporočeni maloprodajni ceni.

Slika 3. 9: Številka ISBN/EAN in črtna koda z dodano petmestno kodo.

3 IBM-ova identifikacijska shema in permutacije

Identifikacijska shema, ki so jo razvili v IBM-u, se uporablja pri kreditni kartici podjetja, krvnih bankah, farmacijah, nemških bankah, ...

Naj $a_1a_2a_3\ldots a_n$, $n \in \mathbb{N}$, predstavlja identifikacijsko število. Zadnja števka a_n je kontrolna številka. Za izračun kontrolne številke uporabimo permutacijo (zapisana v obliki cikla)

$$\sigma = (0)\,(124875)\,(36)\,(9) \in S_{10}$$

na enega izmed naslednjih dveh načinov:

1. n - sodo število:

$$\sigma(a_1) + a_2 + \sigma(a_3) + a_4 + \ldots + \sigma(a_{n-1}) + a_n = 0 \pmod{10},$$

oziroma

$$a_n = (10 - (\sigma(a_1) + a_2 + \sigma(a_3) + a_4 + \ldots + \sigma(a_{n-1})) \bmod 10) \bmod 10,$$

2. n - liho število:

$$a_1 + \sigma(a_2) + a_3 + \sigma(a_4) + \ldots + \sigma(a_{n-1}) + a_n = 0 \pmod{10},$$

oziroma

$$a_n = (10 - (a_1 + \sigma(a_2) + a_3 + \sigma(a_4) + \ldots + \sigma(a_{n-1}))) \bmod 10) \bmod 10.$$

Poglejmo si izračun kontrolne številke na dveh primerih:

1. 00001324136 9 : n = 12

$$\sigma(0) + 0 + \sigma(0) + 0 + \sigma(1) + 3 + \sigma(2) + 4 + \sigma(1) + 3 + \sigma(6) =$$
$$= 0 + 0 + 0 + 0 + 2 + 3 + 4 + 4 + 2 + 3 + 3 = 21 \quad \Rightarrow$$
$$\Rightarrow \quad a_{12} = (10 - 21 \bmod 10) \bmod 10 = 9 \bmod 10 = 9,$$

2. 025003104756 7 : n = 13

$$0 + \sigma(2) + 5 + \sigma(0) + 0 + \sigma(3) + 1 + \sigma(0) + 4 + \sigma(7) + 5 + \sigma(6) =$$
$$= 0 + 4 + 5 + 0 + 0 + 6 + 1 + 0 + 4 + 5 + 5 + 3 = 33 \quad \Rightarrow$$
$$\Rightarrow \quad a_{13} = (10 - 33 \bmod 10) \bmod 10 = 7 \bmod 10 = 7.$$

Pokažimo, da IBM-ova identifikacijska shema zazna vse napake ene števke. Naj bo $a_1a_2a_3\ldots a_n$ identifikacijsko število in n sodo število. Na i-tem mestu, $1 \leq i \leq n$, zamenjamo a_i z a_i', $a_i \neq a_i'$. Označimo $a = \sigma(a_1)+a_2+\sigma(a_3)+a_4+\ldots+\sigma(a_{n-1})+a_n$, a' izračunamo na enak način, le da a_i zamenjamo z a_i'. Odštejmo $a - a'$. Napaka bo odkrita samo v primeru $a - a' \neq 0 \pmod{10}$. Ločimo dva primera:

1. $a - a' = \sigma(a_i) - \sigma(a_i') \neq 0 \pmod{10}$, ker so permutacije injektivne,

2. $a - a' = a_i - a_i' \neq 0 \pmod{10}$.

Podobno lahko pokažemo za n - liho število. Ta shema torej odkrije napako ene števke.

Sedaj pa v številu $a_1a_2a_3 \ldots a_n$, n sodo število, zamenjajmo dve sosednji števki, recimo a_i in a_{i+1}, $1 \leq i \leq n-1$. Izračunamo a in a' (podobno kot zgoraj). Odštejmo $a - a'$. Napaka bo odkrita samo v primeru $a - a' \neq 0 \pmod{10}$. Ločimo dva primera:

1. $a - a' = \sigma(a_i) + a_{i+1} - \sigma(a_{i+1}) - a_i = \sigma(a_i) - a_i - \sigma(a_{i+1}) + a_{i+1}$.

   Razlika bo različna od 0, samo v primeru, ko bo $\sigma(a_i) - a_i \neq \sigma(a_{i+1}) - a_{i+1}$. Če pogledamo permutacijo $\sigma = (0)\,(124875)\,(36)\,(9) \in S_{10}$, vidimo, da bo veljalo $a - a' \neq 0 \pmod{10}$ vedno, razen, če se bosta zamenjali števki 0 in 9.

2. $a - a' = a_i + \sigma(a_{i+1}) - a_{i+1} - \sigma(a_i) = a_i - \sigma(a_i) - a_{i+1} + \sigma(a_{i+1})$.

   Razlika bo različna od 0, samo v primeru, ko bo $a_i - \sigma(a_i) \neq a_{i+1} - \sigma(a_{i+1})$. Če pogledamo permutacijo $\sigma = (0)\,(124875)\,(36)\,(9) \in S_{10}$, vidimo, da bo veljalo $a - a' \neq 0 \pmod{10}$ vedno, razen, če se bosta zamenjali števki 0 in 9.

Podobno lahko pokažemo za n - liho število. Ta shema torej odkrije zamenjavo sosednjih števk ($a_i\, a_{i+1} \rightarrow a_{i+1}\, a_i$), razen, če sta zamenjani števki 0 in 9.

Torej smo že prišli do sheme, v kateri je identifikacijsko število poljubne dolžine sestavljeno iz samih števk. Shema pa odkrije vse napake ene števke in skoraj vse zamenjave sosednjih števk.

4 Verhoeffova identifikacijska shema in teorija grup

Poglejmo si še Verhoeffovo identifikacijsko shemo. Naj $a_1a_2a_3\ldots a_n,\ n\in\mathbb{N}$, predstavlja identifikacijsko število s kontrolno številko a_n. Veljati mora:

$$\sigma^{n-1}(a_1)\circ\sigma^{n-2}(a_2)\circ\sigma^{n-3}(a_3)\circ\ldots\circ\sigma(a_{n-1})\circ a_n=0\,,$$

kjer je

$$\sigma=(0)\,(14)\,(23)\,(56789)\ \in\ S_{10}$$

in $\circ$ - kompozitum permutacij v grupi D_{10}:
$(D_{10},\circ)\ldots$ grupa simetrij pravilnega 5-kotnika $abcde$.

Elementi grupe D_{10}:

$$0=\begin{pmatrix}a&b&c&d&e\\a&b&c&d&e\end{pmatrix},\quad 1=\begin{pmatrix}a&b&c&d&e\\b&c&d&e&a\end{pmatrix},\quad 2=\begin{pmatrix}a&b&c&d&e\\c&d&e&a&b\end{pmatrix},\quad 3=\begin{pmatrix}a&b&c&d&e\\d&e&a&b&c\end{pmatrix},$$

$$4=\begin{pmatrix}a&b&c&d&e\\e&a&b&c&d\end{pmatrix},\quad 5=\begin{pmatrix}a&b&c&d&e\\a&e&d&c&b\end{pmatrix},\quad 6=\begin{pmatrix}a&b&c&d&e\\e&d&c&b&a\end{pmatrix},\quad 7=\begin{pmatrix}a&b&c&d&e\\d&c&b&a&e\end{pmatrix},$$

$$8=\begin{pmatrix}a&b&c&d&e\\c&b&a&e&d\end{pmatrix},\quad 9=\begin{pmatrix}a&b&c&d&e\\b&a&e&d&c\end{pmatrix}$$

in njihovi kompozitumi:

$\circ$	0	1	2	3	4	5	6	7	8	9
0	0	1	2	3	4	5	6	7	8	9
1	1	2	3	4	0	6	7	8	9	5
2	2	3	4	0	1	7	8	9	5	6
3	3	4	0	1	2	8	9	5	6	7
4	4	0	1	2	3	9	5	6	7	8
5	5	9	8	7	6	0	4	3	2	1
6	6	5	9	8	7	1	0	4	3	2
7	7	6	5	9	8	2	1	0	4	3
8	8	7	6	5	9	3	2	1	0	4
9	9	8	7	6	5	4	3	2	1	0

Tabela 4. 1: Kompozitum elementov grupe D_{10}.

Poglejmo si izračun kontrolne številke na dveh primerih:

1. 5701 2: $n = 5$

$$\begin{aligned}
\sigma^4(a_1) \circ \sigma^3(a_2) \circ \sigma^2(a_3) \circ \sigma^1(a_4) \circ a_5 &= 0 \\
\sigma^4(5) \circ \sigma^3(7) \circ \sigma^2(0) \circ \sigma(1) \circ a_5 &= 0 \\
9 \circ 5 \circ 0 \circ 4 \circ a_5 &= 0 \\
(9 \circ 5) \circ (0 \circ 4) \circ a_5 &= 0 \\
4 \circ 4 \circ a_5 &= 0 \\
(4 \circ 4) \circ a_5 &= 0 \\
3 \circ a_5 &= 0
\end{aligned}$$

Pogledamo v tabelo in vidimo, da je edina rešitev $a_5 = 2$.

2. 27163 1: $n = 6$

$$\begin{aligned}
\sigma^5(a_1) \circ \sigma^4(a_2) \circ \sigma^3(a_3) \circ \sigma^2(a_4) \circ \sigma(a_5) \circ a_6 &= 0 \\
\sigma^5(2) \circ \sigma^4(7) \circ \sigma^3(1) \circ \sigma^2(6) \circ \sigma^1(3) \circ a_6 &= 0 \\
3 \circ 6 \circ 4 \circ 8 \circ 2 \circ a_6 &= 0 \\
(3 \circ 6) \circ (4 \circ 8) \circ 2 \circ a_6 &= 0 \\
9 \circ 7 \circ 2 \circ a_6 &= 0 \\
(9 \circ 7) \circ 2 \circ a_6 &= 0 \\
2 \circ 2 \circ a_6 &= 0 \\
(2 \circ 2) \circ a_6 &= 0 \\
4 \circ a_6 &= 0
\end{aligned}$$

Pogledamo v tabelo in vidimo, da je edina rešitev $a_6 = 1$.

Verhoeffova identifikacijska shema odkrije vse napake ene števke, prav tako zamenjavo dveh sosednjih števk. Še več: Verhoeffova identifikacijska shema odkrije vse napake iz tabele 2. 1.

Pokažimo, da Verhoeffova identifikacijska shema zazna vse napake ene števke. Naj bo $a_1a_2a_3 \ldots a_n$ identifikacijsko število. Na i-tem mestu, $1 \leq i \leq n$, zamenjamo a_i z a_i', $a_i \neq a_i'$. Recimo, da napaka ni odkrita. Velja torej:

$$\sigma^{n-1}(a_1) \circ \ldots \circ \sigma^{n-i}(a_i) \circ \ldots \circ a_n = 0$$

in

$$\sigma^{n-1}(a_1) \circ \ldots \circ \sigma^{n-i}(a_i') \circ \ldots \circ a_n = 0 \quad \Rightarrow$$

$$\Rightarrow \quad \sigma^{n-1}(a_1) \circ \ldots \circ \sigma^{n-i}(a_i) \circ \ldots \circ a_n = \sigma^{n-1}(a_1) \circ \ldots \circ \sigma^{n-i}(a_i') \circ \ldots \circ a_n \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \sigma^{n-i}(a_i) = \sigma^{n-i}(a_i')\,.$$

Ker je σ^{n-i} permutacija (bijektivna preslikava) v S_{10}, velja:

$$\sigma^{n-i}(a_i) = \sigma^{n-i}(a_i') \quad \Leftrightarrow \quad a_i = a_i'\,.$$

To pa je v protislovju s predpostavko $a_i \neq a_i'$.

Ta shema torej odkrije napako ene števke.

Leta 1990 je nemška banka (Deutsche Bundesbank) začela uporabljati shemo, ki temelji na Verhoeffovi identifikacijski shemi.

Naj $a_1a_2a_3\ldots a_n,\ n \in \mathbb{N}$, predstavlja identifikacijsko število s kontrolno številko a_n. Veljati mora:

$$\sigma(a_1) \circ \sigma^2(a_2) \circ \sigma^3(a_3) \circ \ldots \circ \sigma(a_{n-1}) \circ a_{n-1} \circ a_n = 0\,,$$

kjer je

$$\sigma = (01589427)\,(36) \ \in \ S_{10}$$

in $\circ$ - kompozitum permutacij v grupi D_{10}:
$(D_{10}, \circ)\ldots$ grupa simetrij pravilnega 5-kotnika *abcde*.

Poglejmo si serijske številke na bivših nemških markah. Številke so enajstmestne, sestavljene iz števk in črk.

Slika 4. 1: Bankovec za 10 nemških mark s serijsko številko GK7042314S in kontrolno številko 5.

Da lahko uporabimo shemo, moramo najprej vse črke spremeniti v številke. To naredimo s pomočjo naslednje tabele:

A	D	G	K	L	N	S	U	Y	Z
0	1	2	3	4	5	6	7	8	9

Tabela 4. 1: Numerične vrednosti za črke uporabljene na bankovcih.

Poglejmo si izračun kontrolne številke na primeru s slike 4. 1 : GK7042314S 5 . Najprej spremenimo vse črke v številke: 2370423146 5.

$$\begin{aligned}
\sigma(a_1) \circ \sigma^2(a_2) \circ \sigma^3(a_3) \circ \sigma^4(a_4) \circ \sigma^5(a_5) \circ \sigma^6(a_6) \circ \sigma^7(a_7) \circ \\
\circ \sigma^8(a_8) \circ \sigma^9(a_9) \circ \sigma^{10}(a_{10}) \circ a_{11} = 0 \\
\sigma(2) \circ \sigma^2(3) \circ \sigma^3(7) \circ \sigma^4(0) \circ \sigma^5(4) \circ \sigma^6(2) \circ \sigma^7(3) \circ \\
\circ \sigma^8(1) \circ \sigma^9(4) \circ \sigma^{10}(6) \circ a_{11} = 0 \\
7 \circ 3 \circ 5 \circ 9 \circ 5 \circ 9 \circ 6 \circ 1 \circ 2 \circ 6 \circ a_{11} = 0 \\
(7 \circ 3) \circ (5 \circ 9) \circ (5 \circ 9) \circ (6 \circ 1) \circ (2 \circ 6) \circ a_{11} = 0 \\
9 \circ 1 \circ 1 \circ 5 \circ 8 \circ a_{11} = 0 \\
(9 \circ 1) \circ (1 \circ 5) \circ 8 \circ a_{11} = 0 \\
8 \circ 6 \circ 8 \circ a_{11} = 0 \\
(8 \circ 6) \circ 8 \circ a_{11} = 0 \\
2 \circ 8 \circ a_{11} = 0 \\
(2 \circ 8) \circ a_{11} = 0 \\
5 \circ a_{11} = 0
\end{aligned}$$

Pogledamo v tabelo in vidimo, da je edina rešitev $a_{11} = 5$.

Verhoeffova identifikacijska shema je od vseh omenjenih najboljša, saj odkrije vse napake iz tabele 2. 1 in jo lahko uporabimo za identifikacijsko število poljubne dolžine.

Na tem mestu je utemeljeno vprašanje: Zakaj večina identifikacijskih števil uporablja shemo kontrolne številke, ki ne odkrije vseh napak iz tabele 2. 1 ? To vprašanje ostane brez konkretnega odgovora tudi v [2] in [4]. V [4] sicer predvidevajo, da je zato krivo zgodovinsko naključje in pomanjkanje znanja o obstoječih metodah.

5 Primeri identifikacijskih shem v Sloveniji

5.1 EMŠO

Osebna identifikacijska številka v Sloveniji je trinajstmestna enotna matična številka občana (EMŠO), ki jo določa ministrstvo, pristojno za notranje zadeve kot upravljalec Centralnega registra prebivalstva (CRP) po algoritmu, ki vključuje modul 11. Določi ga na osnovi sedemmestnega datuma rojstva in spola. Številke od prvega do sedmega mesta EMŠO - ja so datum rojstva - 2 - mestni dan, 2 - mestni mesec in 3 - mestna letnica (npr. datum rojstva 2. 5. 1982 zapišemo 0205982). Na osmem in devetem mestu je številka registra - šifra 50. Na desetem, enajstem in dvanajstem mestu je zaporedna številka, na trinajstem mestu pa kontrolna številka. Zaporedna številka je kombinacija spola in zaporedne številke za osebe, rojene istega dne - 000-499 za moške in 500-999 za ženske. Osnova za izračun zaporedne številke je evidenca določenih EMŠO - jev: zaporedni številki zadnjega EMŠO - ja, ki je bil določen na isti datum rojstva in spol, se prišteje vrednosti 1. Tako dobljeni seštevek je zaporedna številka za EMŠO. Na osnovi prvih dvanajstih številk se izračuna kontrolna številka. V evidenci določenih EMŠO - jev nekatere zaporedne in kontrolne številke pri posameznem datumu rojstva niso možne.

Shema kontrolne številke za EMŠO $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}$ je naslednja:

$$7\,a_1+6\,a_2+5\,a_3+4\,a_4+3\,a_5+2\,a_6+7\,a_7+6\,a_8+5\,a_9+4\,a_{10}+3\,a_{11}+2\,a_{12}+a_{13} = 0 \pmod{11}$$

oziroma

$$(7,6,5,4,3,2,7,6,5,4,3,2,1)\cdot(a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11},a_{12},a_{13}) = 0 \pmod{11}.$$

Kontrolno številko lahko torej izračunamo po formuli:

$$a_{13} = (11 - (7\,a_1 + 6\,a_2 + 5\,a_3 + 4\,a_4 + 3\,a_5 + 2\,a_6 + 7\,a_7 + \\ + 6\,a_8 + 5\,a_9 + 4\,a_{10} + 3\,a_{11} + 2\,a_{12}) \bmod 11) \bmod 11.$$

Če dobimo $a_{13} = 10$, to ni veljavna kontrolna številka, ker mora biti kontrolna številka enomestna. V takem primeru upravljalec CRP zaporedno številko preskoči, vrednost zaporedne številke poveča za 1, izračun kontrolne številke pa ponovi po istem postopku.

Poglejmo si izračun kontrolne številke na primeru: 1804978505088

$$7\cdot 1+6\cdot 8+5\cdot 0+4\cdot 4+3\cdot 9+2\cdot 7+7\cdot 8+6\cdot 5+5\cdot 0+4\cdot 5+3\cdot 0+2\cdot 8 = 234 \quad \Rightarrow$$

$$\Rightarrow \quad a_{13} = (\,11 - 234 \bmod 11\,) \bmod 11 = 8 \bmod 11 = 8$$

5.2 Osebna izkaznica in potni list

Številka osebne izkaznice in številka potnega lista sta določeni avtomatično po zaporedju števil od 1 dalje. Za številčenje osebnih izkaznic je predvidenih 9 mest. Prva osebna izkaznica je imela številko 000000001. Pri številčenju potnega lista pa je poleg črkovne oznake, ki označuje vrsto potnega lista (P je npr. oznaka za običajni potni list, D za diplomatskega, S za službenega itd.) na razpolago še 8 mest za števila. Prvi potni list je imel številko P00000001. Številka osebne izkaznice in številka potnega lista sta torej brez kontrolne številke.

Poglejmo pa is standardiziran zapis podatkov na območju strojno čitljive kode (v OCR - B formatu) na osebni izkaznici in potnem listu, ki je namenjen branju potnih listin s pomočjo čitalcev OCR - B zapisa na mejnih prehodih in omogoča hitrejše prehajanje državne meje.

Na OCR - B zapisu na osebni izkaznici imamo 4 kontrolne številke, na potnem listu pa 5. Če je pri računanju kontrolne številke vključena črka, ji priredimo numerično vrednost po tabeli 6. 1, znaku $<$ pa vrednost 0.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35

Tabela 6. 1: Numerične vrednosti za črke uporabljene na OCR - B zapisu.

Kontrolne številke na OCR - B zapisu se izračunajo na naslednji način: recimo, da računamo kontrolno številko a_n, $n \in \mathbb{N}$, za identifikacijsko število $a_1 a_2 a_3 \ldots a_{n-1}$, potem je

$$a_n = (7 \cdot a_1 + 3 \cdot a_2 + 1 \cdot a_3 + 7 \cdot a_4 + 3 \cdot a_5 + 1 \cdot a_6 + \ldots + k \cdot a_{n-1}) \bmod 10$$

oziroma

$$a_n = (\,(7, 3, 1, 7, 3, 1, 7, 3, 1, \ldots, k) \cdot (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, \ldots a_{n-1})\,) \bmod 10\,,$$

pri čemer je število $k \in \{7, 3, 1\}$ odvisno od n.

Osebna izkaznica:

OCR - B zapis na osebni izkaznici je na njeni hrbtni strani in obsega tri vrstice. Pri obliki in dimenziji natisnjene podobe v OCR - B formatu je upoštevan mednarodni standard ISO 1073/II:1976 - Alphanumeric character sets for optical recognition - Part 2: Character set OCR - B - shapes and dimensions of the printed image ter ISO 1831:1980 - Printing specifications for optical character recognition (OCR - B zapis).

1. Prva vrstica ima prvih 15 mest predpisanih:

- Prvi dve mesti sta namenjeni za oznako dokumenta (obvezne so prve črke: A, V ali I). Slovenska osebna izkaznica ima oznako **I<** (pri osebni izkaznici je I obvezna - Identity Card) .
- Sledijo trije znaki za oznako države ali organizacije, ki je izdala izkaznico. Pri nas je to zapis **SI<**.
- Tem petim znakom sledi **devetmestna serijska številka** in **prva kontrolna številka**, izračunana iz devetemstne serijske številke.
- Od šestnajste do tridesete pozicije je prostor za poljubni, nacionalni izbor zapisa. Na naši osebni izkaznici je to **zapis 13 mestnega EMŠO - ja**, zadnja dva znaka sta <<.

2. Druga vrstica ima predpisani dve podatkovni polji in tri kontrolne izračune:
   - Na prvih šestih mestih je zapisan **datum rojstva** imetnika osebne izkaznice v formatu **LLMMDD**; na sedmi poziciji je **kontrolna številka**, izračunana iz teh šest številk.
   - Na osmi poziciji je **oznaka spola**: F = ženski, M = moški, X = nedoločen.
   - Od devete do štirinajste pozicije je zapisan **datum veljavnosti** v formatu **LLMMDD**, na petnajsti poziciji je **kontrolna številka**, izračunana iz teh šest številk.
   - Sledi oznaka države: **SI** in nato do vključno 29 znaka <.
   - Trideseta pozicija je **kontrolna številka** za vsa polja in kontrolna številka prvih dveh vrstic (v prvi vrstici izpustimo prvih pet mest, v drugi vrstici izpustimo oznako spola, vsem preostalim znakom priredimo numerične vrednosti in nato izračunamo kontrolno številko).

3. Tretja vrstica ima poljuben zapis, praviloma pa se uporablja za zapis priimka in imena. Sintaksa zapisa je naslednja:

   priimek<<ime, če pa ima posameznik priimek ali ime sestavljen iz dveh besed, sta ti medsebojno ločeni z znakom <. V primeru, da dolžina priimka in imena presega 30 znakov, se izpiše dominantno ime oziroma priimek ali primerna okrajšava po imetnikovi izbiri. Če ustrezne okrajšave priimka ali imena ni, se podatki izpišejo po vrsti, dokler gre (prvih trideset znakov). Za zapisovanje imen in priimkov je uporabljena transkripcija 1/2: Č = CH, Š = SS, Ž = ZZ, Ć = CC, Đ = DJ, DŽ = DZ, Ä = AE, Ö = OE, Ü = UE.

Slika 6. 1: OCR - B zapis na osebni izkaznici.

Potni list:

Strojno čitljiv zapis je izdelan v skladu z mednarodnim standardom ICAO Doc 9303 - Machine Readable Travel Documents, Machine Readable Passport (Fourth Edition - 1999). Je dvovrstičen in vključuje naslednje podatke:

1. Prva vrstica:
   - Prvi dve podatkovni polji označujeta tip dokumenta in se zapišeta kot **P<** (običajni potni list), kot **PD** (diplomatski potni list) ali kot **PS** (službeni potni list).
   - Sledeča tri podatkovna polja so rezervirana za vpis kode države, ki je izdala potni list. Za Republiko Slovenijo je to zapis **SVN**.
   - 6 do 44 polje (39 polj) je namenjeno osebnemu imenu imetnika potnega lista. Zapiše se kot **priimek<<ime**. Med dvema priimkoma oz. dvema imenoma je <. Transkripcija zapisa imena in priimka na potnem listu je 1/1 z izjemami pri latinskih črkah: Ä, Ĺ, Ń, Ö, Ř, Ü, kjer velja pravilo 1/2, kar je skladno s standardom ICAO Doc 9303.
2. Druga vrstica:
   - Prvih devet podatkovnih polj je namenjeno zapisu številke potnega lista na način **P12345678** (običajni potni list), **D12345678** (diplomatski potni list) oziroma **S12345678** (službeni potni list).
   - 10 podatkovno polje je **kontrolna številka**, izračunana iz prvih devetih mest (črki priredimo numerično vrednost).
   - Naslednja tri polja so namenjena vpisu državljanstva imetnika potnega lista na način **SVN**.
   - Od polja 14 do 19 je zapisan **rojstni datum** imetnika na način **LLMMDD**.
   - 20 mesto je namenjeno **kontrolni številki**, ki jo izračunamo iz zadnjih šestih števk (datum rojstva).
   - Na 21 poziciji je **oznaka spola**: F = ženski, M = moški, < = nedoločen.
   - Od 22 do 27 pozicije je zapisan **datum veljavnosti** v formatu **LLMMDD**.
   - Na 28 mestu je **kontrolna številka**, izračunana iz zadnjih šestih števk (datum veljavnosti).
   - Od 29 do 42 mesta je zapis **EMŠO** - ja, na 42 mestu je <.
   - 43 pozicija je določena za zapis **kontrolne številke**, ki jo izračunamo iz EMŠO - ja.
   - 44 pozicija pa je določena za zapis **kontrolne številke** cele druge vrstice (v drugi vrstici izpustimo zapis državljanstva (**SVN**) in oznako spola, vsem preostalim znakom priredimo numerične vrednosti in nato izračunamo kontrolno številko).

Slika 6. 2: OCR - B zapis na potnem listu.

Poglejmo si izračun kontrolne številke na primeru prvih devetih mest druge vrstice na OCR - B zapisu na potnem listu: P00755169. Najprej moramo črki P prirediti numerično vrednost po tabeli 6. 1: črka P ima vrednost 25.

$$\begin{aligned} a_{10} &= (\,(7,3,1,7,3,1,7,3,1)\cdot(a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9)\,) \bmod 10 = \\ &= (7\cdot a_1 + 3\cdot a_2 + 1\cdot a_3 + 7\cdot a_4 + 3\cdot a_5 + 1\cdot a_6 + 7\cdot a_7 + 3\cdot a_8 + 1\cdot a_9) \bmod 10 = \\ &= (7\cdot 25 + 3\cdot 0 + 1\cdot 0 + 7\cdot 7 + 3\cdot 5 + 1\cdot 5 + 7\cdot 1 + 3\cdot 6 + 1\cdot 9) \bmod 10 = \\ &= 278 \bmod 10 = \\ &= 8. \end{aligned}$$

5.3 Kartica zdravstvenega zavarovanja

Na kartici zdravstvenega zavarovanja imamo dve številki: številka izdajatelja in številka zdravstvenega zavarovanja.

Številka izdajatelja je enajstmestna in za vse zavarovance obveznega zdravstvenega zavarovanja enaka. Sestavljena je iz:

- šifre panoge (2 mesti) **80**,
- kode države (3 mesta) **705**,
- identifikacije izdajatelja (5 mest) **00001** in kontrolne številke, ki se izračuna po Luhnovi formuli izračunano po modulu 10.

Način izračuna kontrolne številke:

$$(1,2,1,2,1,2,1,2,1,2,1)\cdot(a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11}) = 0 \pmod{10}$$

oziroma

$$\sum_{i=2k-1,\,k\in\mathbb{N}}^{11} a_i + 2\cdot\sum_{i=2k,\,k\in\mathbb{N}}^{10} a_i = 0 \pmod{10}\,.$$

Kontrolno številko lahko torej izračunamo po formuli:

$$a_{11} = \left(10 - \left(\sum_{i=2k-1,\,k\in\mathbb{N}}^{9} a_i + 2\cdot\sum_{i=2k,\,k\in\mathbb{N}}^{10} a_i\right) \bmod 10\right) \bmod 10\,.$$

Primer: 8070500001 8

$$\begin{aligned} a_{11} &= (10 - (\,(1,2,1,2,1,2,1,2,1,2)\cdot(8,0,7,0,5,0,0,0,0,1\,) \bmod 10) \bmod 10 = \\ &= (10 - (8 + 2\cdot 0 + 7 + 2\cdot 0 + 5 + 2\cdot 0 + 0 + 2\cdot 0 + 0 + 2\cdot 1) \bmod 10) \bmod 10 = \\ &= (10 - 22 \bmod 10) \bmod 10 = \\ &= 8 \bmod 10 = \\ &= 8. \end{aligned}$$

Številka zdravstvenega zavarovanja je devetmestna. Sestavljena je iz tekoče številke, ki je večja od 2.000.000 in kontrolne številke.

Način izračuna kontrolne številke:

$$a_9 = (11 - (6\cdot a_1 + 7\cdot a_2 + 2\cdot a_3 + 3\cdot a_4 + 4\cdot a_5 + 5\cdot a_6 + 6\cdot a_7 + 7\cdot a_8) \bmod 11) \bmod 11\,.$$

Če dobimo $a_9 = 10 \quad \Rightarrow \quad a_9 = 0\,.$

Primer: 03957246 7

$$\begin{aligned} a_9 &= (11 - (\,(6,7,2,3,4,5,6,7) \cdot (0,3,9,5,7,2,4,6)\,) \bmod 11) \bmod 11 = \\ &= (11 - (6 \cdot 0 + 7 \cdot 3 + 2 \cdot 9 + 3 \cdot 5 + 4 \cdot 7 + 5 \cdot 2 + 6 \cdot 4 + 7 \cdot 6) \bmod 11) \bmod 11 = \\ &= (11 - 158 \bmod 11) \bmod 11 = \\ &= 7 \bmod 11 = \\ &= 7. \end{aligned}$$

5.4 Številka transakcijskega računa in serijske številke na bankovcih

Vsi transakcijski računi in njihovi imetniki so evidentirani v Registru transakcijskih računov, ki ga upravlja Banka Slovenije, ažurni podatki iz registra pa so dnevno na razpolago bankam, Davčni upravi RS, Carinski upravi RS, Upravi RS za javna plačila, Agenciji RS za javnopravne evidence in storitve ter sodiščem. Podatki o transakcijskih računih in njihovih imetnikih so javno dostopni, razen transakcijskih računov fizičnih oseb - posameznikov, ki so zaradi zakona o varstvu osebnih podatkov skriti javnosti. Za učinkovito delovanje registra je nujno potrebno, da imajo enotno strukturo vsi transakcijski računi, ki so v njem evidentirani. Po vsebini transkacijski račun ni enoten. Zakonodaja, pa tudi drugi, včasih čisto praktični dejavniki so podlaga razdelitvi TRR na različne "vrste računa".

Z razliko od svojih predhodnikov (tekočega oz. žiro računa) se uporablja za določitev strukture transakcijskega računa enoten standard. Med prednostmi enotne strukture transakcijskega računa je lažje uvajanje elektronskega poslovanja. Enotna struktura računa olajša optično branje podatkov, zapis podatkov na magnetne medije in prenos podatkov na elektronski način. Banka lahko sredstva, ki jih prejme za svojo stranko na svoj transakcijski račun, takoj avtomatično usmeri na transakcijski račun stranke, ne da bi morala opraviti kakršen koli ročni poseg v dokumentacijo. Postopki v elektronski izmenjavi podatkov brez ročne intervencije so definirani kot "Straight through processing" (STP) ter zagotavljajo pravilno in hitro usmeritev plačila od plačnika do njegovega poslovnega partnerja.

Nacionalna struktura transakcijskega računa v Evropi

Države Evropske unije (European Union - EU) in Evropskega ekonomskega prostora (European economic Area - EEA) si že od začetka svojega povezovanja prizadevajo, da bi poenotile in s tem pocenile medsebojno komuniciranje na vseh področjih gospodarstva in tudi na področju plačilnega prometa. Evropski komite za bančne standarde (European Committee for Banking Standards - ECBS), je na osnovi analize vseh struktur domačih bančnih računov v državah EU in EEA pripravil predloge za poenotenje strukture bančnega transakcijskega računa. Bančni računi po posameznih državah imajo tako imenovano "domaco" osnovno številko bančnega računa (Basic Bank Account Number - **BBAN**). BBAN so v EU strukturirani različno glede na nacionalne standarde, imajo različno dolžino znakov, ki so lahko črkovni in numerični, uporabljajo različne načine kontrole itd. Izkazalo se je tudi, da imajo vsi ti BBAN določene skupne značilnosti. V vsaki strukturi se pojavlja:

- identifikacija banke oz. njene podružnice (b),
- identifikacija bančne stranke (a) in
- kontrolna številka (c).

Pretežno je BBAN v Evropi sestavljen na dva načina: $\mathbf{b} + \mathbf{a} + \mathbf{c}$ ali $\mathbf{a} + \mathbf{c} + \mathbf{b}$.

Mednarodna struktura transakcijskega računa - IBAN

Na osnovi analize nacionalnih številk bančnih računov je ECBS v letu 1996 pripravil predlog t.i. mednarodne številke bančnega računa (International bank account number - **IBAN**). IBAN poleg banke in stranke identificira tudi državo in vsebuje nadaljnje kontrole pravilnosti celotne številke računa. Z uveljavitvijo IBAN je omogočena avtomatična obdelava mednarodnih transakcij. V publikaciji Register of European Account Numbers (TR201 V2.2.21) na spletni strani ECBS so objavljene domače in mednarodne številke bančnih računov v Evropi. Od marca 2003 je v tem seznamu tudi Slovenija.
IBAN je sestavljen tako, da se pred BBAN postavijo štirje znaki:

- dva črkovna znaka - oznaki za državo po ISO standardu (d) in
- dva numerična kontrolna znaka (e).

Celotni račun v strukturi IBAN je sestavljen takole: **IBAN = d + e + BBAN**.

Struktura transakcijskega računa v Sloveniji

V Sloveniji je bila enotna struktura transakcijskega računa (BBAN) na podlagi priporočil ECBS potrjena že v letu 1995, uvajala pa se je postopoma hkrati z reformo plačilnih sistemov (uvedba sistema BPRC in sistema žiro kliringa) ter prenosom plačilnega prometa za pravne osebe v bančno okolje. S preoblikovanjem tekočih in žiro računov občanov v strukturo transakcijskega računa je bil proces zaključen in transakcijski računi vseh bančnih strank imajo ne glede na status stranke enako strukturo. **Petmestna identifikacija banke (b)** se deli na **dvomestno identifikacijo same banke (b1)** in na **trimestno identifikacijo njene organizacijske enote (b2)**. Informacija o dvomestnih identifikacijskih številkah bank je dostopna na naslovu: Podatki o bankah, hranilnicah in drugih financnih institucijah. Šifrant vseh organizacijskih enot bank ni posebej objavljen, vir tega podatka je le vsaka posamezna banka za svoje enote. Določitev **osemmestne identifikacije komitenta (a)** je prepuščena vsaki banki posebej. **Dvomestna kontrolna številka (c)** se izračuna po poenostavljenem postopku standarda ISO 7064, MOD 97-10.

Slovenski transakcijski račun (BBAN) je sestavljen po sistemu **b + a + c**. Vsi znaki so numerični.

IBAN za transakcijske račune v Sloveniji se zapiše tako, da se pred transakcijski račun postavi oznaka Slovenije po ISO standardu in dva numerična kontrolna znaka:

- d: oznaka Slovenije - 2 mesti (**SI**),
- e: kontrolna številka - 2 mesti (izračun na podlagi standarda ISO 13616).

V registru transakcijskih računov podatki o številki transakcijskega računa vsebujejo tudi podatek o IBAN.

Primer: elektronski zapis oz. izpis IBAN komitenta Abanke: SI56051008000032875 (brez presledkov oz. ločilnih znakov).

Pa si poglejmo še shemo kontrolne številke:

- **izračun kontrolne številke v strukturi transakcijskega računa (BBAN)**:
  v številki transakcijskega računa $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}a_{14}a_{15}$ je $a_{14}a_{15}$ kontrolna številka; velja:

$$a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}a_{14}a_{15} = 1 \pmod{97}$$

  oziroma

$$a_{14}a_{15} = 98 - (a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}00) \bmod 97\,.$$

  Izračunajmo kontrolno številko transakcijskega računa na prej omenjenem primeru: 051008000032875

$$a_{14}a_{15} = 98 - 051008000032800 \bmod 97 = 98 - 23 = 75\,.$$

- **izračun kontrolne številke IBAN**:
  kontrolna številka za transakcijski račun v strukturi IBAN je določena po mednarodnem standardu ISO 13616, ki je izšel leta 1997 v publikaciji Banking and Related Financial Services - International Bank Account Number (IBAN). Najprej v strukturi IBAN = d + e + BBAN spremenimo vrstni red v IBAN = BBAN + d + e. Namesto dveh črk v "d" (oznaki za državo) zapišemo numerično vrednost po tabeli 6. 1. Tako dobimo 21 - mestno število. Prejšnji primer:

$$SI56051008000032875 \quad \rightarrow \quad 051008000032875281856\,.$$

  Veljati mora:

$$a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}a_{14}a_{15}a_{16}a_{17}a_{18}a_{19}a_{20}a_{21} = 1 \pmod{97}\,.$$

  Izračun kontrolne številke "e" je torej naslednji:

$$a_{20}a_{21} = 98 - (a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{13}a_{14}a_{15}a_{16}a_{17}a_{18}a_{19}00) \bmod 97\,.$$

  Izračunajmo kontrolno številko IBAN na zgornjem primeru: SI56051008000032875

$$a_{20}a_{21} = 98 - 051008000032875281800 \bmod 97 = 98 - 42 = 56\,.$$

Serijske številke na bankovcih Banke Slovenije

Serijske številke na bankovcih Banke Slovenije so sestavljene iz dveh črk, ki so izbrane po slučajnostnem izboru in sedmih mest za številke pri bankovcu za 10.000 SIT, oziroma šestih mest za številke pri vseh ostalih bankovcih. Številke potekajo zaporedno, od 000001 do 1000000 (zadnja serijska številka je pri bankovcih izjemoma sedem mestna) oziroma 9999999 (za 10.000 SIT), v odvisnosti od velikosti naročila tiskanja. Serijske številke za nadomestno serijo bankovcev tj. bankovci, ki nadomeščajo v procesu izdelave zaradi slabe kakovosti izločene bankovce, pa imajo črkovno kombinacijo AZ oziroma ZA in zaporedno številčenje.

Serijska številka na bankovcih Banke Slovenije je brez kontrolne številke.

5.5 Davčna številka

Zakon o davčni službi (ZDS, 30. člen) opredeljuje davčno številko kot identifikacijski znak, ki označuje davčnega zavezanca in se uporablja za enotno opredelitev in povezavo podatkov v davčnih evidencah, ki jih davčni organ vodi o davčnem zavezancu.

Davčna številka je osemmestna številka $a_1a_2a_3a_4a_5a_6a_7a_8$. Prvih sedem mest je osnovna številka, ki je naključno izbrana iz nabora številk od 1000000 do 9999999. Na osmem mestu je kontrolna številka. Velja:

$$(8,7,6,5,4,3,2,1) \cdot (a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8) = 0 \pmod{11}$$

oziroma

$$8 \cdot a_1 + 7 \cdot a_2 + 6 \cdot a_3 + 5 \cdot a_4 + 4 \cdot a_5 + 3 \cdot a_6 + 2 \cdot a_7 + 1 \cdot a_8 = 0 \pmod{11}$$

in

$$a_8 = (11 - (8 \cdot a_1 + 7 \cdot a_2 + 6 \cdot a_3 + 5 \cdot a_4 + 4 \cdot a_5 + 3 \cdot a_6 + 2 \cdot a_7) \bmod 11) \bmod 11\,.$$

Če dobimo $a_8 = 0$, se ta osnovna številka izključi iz nabora možnih davčnih številk. Če je $a_8 = 10 \quad \Rightarrow \quad a_8 = 0$.

Poglejmo si izračun kontrolne številke še na primeru: 22568468

$$\begin{aligned} a_8 &= (11 - (\,(8,7,6,5,4,3,2) \cdot (2,2,5,6,8,4,6)\,) \bmod 11) \bmod 11 = \\ &= (11 - (8 \cdot 2 + 7 \cdot 2 + 6 \cdot 5 + 5 \cdot 6 + 4 \cdot 8 + 3 \cdot 4 + 2 \cdot 6) \bmod 11) \bmod 11 = \\ &= (11 - 146 \bmod 11) \bmod 11 = \\ &= 8 \bmod 11 = \\ &= 8. \end{aligned}$$

5.6 Šifre na maturi

Pojdimo z zadnjim primerom še malce v šolstvo. Na maturi srečamo naslednje šifre:

- **šifre kandidatov** so šestmestne $a_1a_2a_3a_4a_5a_6$. Prvih 5 mest je naključna številka, zadnja pa je kontrolna atevilka:

  $$a_6 = (\,(1,2,3,4,5)\cdot(a_1,a_2,a_3,a_4,a_5)\,) \bmod 9\,,$$

  oziroma

  $$a_6 = (1\cdot a_1 + 2\cdot a_2 + 3\cdot a_3 + 4\cdot a_4 + 5\cdot a_5) \bmod 9\,.$$

  Kandidati dobijo šifre v obliki nalepk s črtno kodo (Code 2of5).

- **šifre ocenjevalcev** so petmestne številke, ki so zaporedne številke v podatkovni bazi ocenjevalcev. So brez kontrolne številke. Ocenjevalci dobijo šifre v obliki nalepk s črtno kodo (Code 93).

- **šifre izpitne pole** so sestavljene iz:
  - vrsta preverjanja (npr. **M** - splošna matura, **P** - poklicna matura),
  - šifre roka (npr. 041 = 2004 - spomladanski rok),
  - šifre predmeta (3-mesta),
  - številka termina znotraj izpitnega roka (1 ali 2),
  - številka gradiva (1 = prva izpitna pola, 2 = druga izpitnapola ...),
  - prevodi gradiv imajo še na zadnjem mestu **I** (italijanski prevod) ali **M** (madžarski).

  Šifre gradiv so brez kontrolne številke in so prav tako natisnjene na poli v obliki črtne kode (Code 39).

  Primer: M041-401-1-1 (splošna matura, spomladanski rok 2004, matematika, termin 1, prva izpitna pola).

Dodatek

Za konec si poglejmo še nekaj malega o črtnih kodah.

> **Črtna koda** je zakodiran zapis identifikacijske številke v obliki različno širokih črtic in presledkov, ki omogoča avtomatsko branje.

Črtna koda je torej sestavljena iz zaporedja (navpičnih) črt in vmesnih presledkov. Črte in presledki so različnih širin in tako jih lahko razumemo kot različne številke ali črke, odvisno od dogovora. Recimo, da takšno kodo preberemo s svetlobnim peresom. Žarek peresa pomaknemo preko črtne kode, pero preko odbitega žarka razbere vzorec ožjih in širših črt ter vmesnih presledkih in ga pošlje procesorju, ki ta vzorec spremeni v zaporedje številk. Pri tem je pomembno, da svetlobno pero pomaknemo preko črtne kode kar se da enakomerno hitro, saj bi na primer krajša upočasnitev za prejeti signal v peresu pomenila isto, kot da je trenutno osvetljena črta ali presledek malo širši, kot je v resnici. Zaradi možnosti napake pri branju črtne kode so danes na voljo vse boljše in boljše, a tudi dražje, priprave za branje. Na primer pri laserski pripravi s pomičnim žarkom laserski žarek večkrat (tudi po stokrat v eni sekundi) prečeše črtno kodo in tako lahko s primerjavo rezultatov dosežemo večjo natančnost in manjšo možnost napake. Drugi način, ki tudi zmanjšuje možnost napake pri branju, pa je, kot smo pokazali v prejšnjih poglavjih, skrit v sami kodi[16] (kontrolne številke).

Povedali smo že, da se v sistemu EAN.UCC uporabljajo tri različne simbologije črtne kode:

- EAN/UPC: EAN-13, UPC-A ali EAN-8, UPC-E,
- ITF-14,
- UCC.EAN-128.

Ko se odloča med različnimi simbologijami in namestitvijo, se upošteva naslednje:

- prostor, ki je na voljo na artiklu,
- vrsta informacije, ki jih je treba zakodirati,
- delovno okolje, v katerem se odčitava simbol črtne kod.

Pa si poglejmo zgradbo črtne kode EAN-13.

[16] To lahko opazimo v trgovini, kjer mora prodajalka včasih tudi dvakrat ali trikrat prebrati isto kodo.

Zgradba črtne kode EAN - 13

Črtna koda EAN - 13 je zgrajena na naslednji način [1]:

- za določitev začetka, konca in sredine črtne kode imamo tri pare (nekoliko daljših) navpičnih črt, ki služijo le kontroli pri branju,

Slika 1: Črtna koda EAN - 13.

- širina presledka med črtama v vsakem od teh parov določa osnovno širino h kode (isto širino imata tudi obe črti),
- vmesni prostor je razdeljen na ustrezno število intervalov širine $7h$, vsak interval pa na 7 pasov širine h,
- če je tak pas bel, to pomeni bit 0, črn pas pa ustreza bitu 1,
- vsaka števka ima 7-bitno kodo, ki je odvisna od tega, na katerem mestu se nahaja,
- števke iz prve polovice so kodirane tako, kot kažeta prvi in drugi stolpec, števke iz druge polovice črtne kode pa tako, kot prikazuje tretji stolpec tabele 1,

Števka	Levi del - koda A	Levi del - koda B	Desna polovica
0	0001101	0100111	1110010
1	0011001	0110011	1100110
2	0010011	0011011	1101100
3	0111101	0100001	1000010
4	0100011	0011101	1011100
5	0110001	0111001	1001110
6	0101111	0000101	1010000
7	0111011	0010001	1000100
8	0110111	0001001	1001000
9	0001011	0010111	1110100

Tabela 1: Bitno kodiranje v črtni kodi EAN - 13.

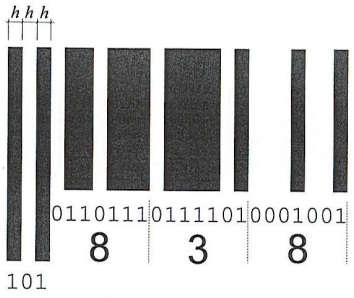
- za števke iz prve polovice ni natanko določeno, kdaj uporabimo kodo A in kdaj kodo B.

Podroben pregled črtne kode s slike 1, katere začetni del je močno povečan na sliki 2, pokaže, da črtna koda ustreza zaporedju bitov:

$$\underbrace{101}_{A}\,\underbrace{0110111}_{8}\,\underbrace{0111101}_{3}\,\underbrace{0001001}_{8}\,\underbrace{0001001}_{8}\,\underbrace{0100111}_{0}\,\underbrace{0001101}_{0}\,\underbrace{01010}_{B}$$

$$\underbrace{1110010}_{0}\,\underbrace{1110010}_{0}\,\underbrace{1110010}_{0}\,\underbrace{1000100}_{7}\,\underbrace{1001110}_{5}\,\underbrace{1010000}_{6}\,\underbrace{101}_{C}$$

Pri tem zaporedje bitov A, B in C ustrezajo levemu robu, sredini in desnemu robu.



Slika 2: Povečava začetka črtne kode.

Ker iz sedmih lahko sestavimo kar 128 različnih kod, v EAN - 13 pa je porabljenih le deset izmed njih, nam takšno kodiranje omogoča dodatno preverjanje napak. Omenimo še, da se iz opisa zgradbe črtne kode vidi, da čitalnik črtne kode ugotovi, ali bere z leve proti desni ali z desne proti levi.

Primeri enodimenzionalnih črtnih kod

Slika je iz [3]:

TABLE 4.4 Features of leading 1D barcodes

GTIN-14		Numeric	Yes	14 13 + CS	Retail POS EAN.UCC (global)
EAN-13		Numeric	Yes	13 12 + CS	Retail POS EAN.UCC (outside US and Canada)
UPC A		Numeric	Yes	12 11 + CS	Retail POS EAN.UCC (US and Canada)
ISBN Bookland with extension		Numeric	Yes	13 12 + CS Plus 2 + CS	Retail POS ISBN, ISSN, ISMN (global)
ITF-14 Interleaved 2-of-5		Numeric	Yes	14 13 + CS	Shipping EAN.UCC
Interleaved 2-of-5 Airline		Numeric		10/12 code identifies inbound carrier and unique ID Even number	Airline bags IATA
Interleaved 2-of-5 Industrial		Numeric	Optional	Variable Even number	EAN.UCC, HIBCC, USPS special services, Optical Product Code Council CIP HR (France), Deutsche Post Identcode and Leitcode
Code 128		Full 128 ASCII + control	Yes	Variable	Airline cargo (IATA) HIBCC USPS delivery confirmation SISAC
EAN.UCC-128		Full 128 ASCII + control	Yes	Variable	Shipping EAN.UCC Containers containing EAN.UCC code products US FDA
Codabar		Numeric plus $ - : / . +	None defined	Variable	Blood banks, libraries US blood centres Ameritech library services
Code 39		Uppercase A–Z, 0–9, space - . $ / + %	Optional	Variable	Wide use for many applications in industry: auto (AIAG), commerce, military (DoD), health (HIBCC)
Code 93		128 ASCII	Yes	Variable	Industry

Slika 3: Enodimenzionalne črtne kode.

Literatura

[1] B. Mohar: *O črtnih kodah*, Presek, **24** (1996-97) 20 – 24

[2] J. Kirtland: *Identification Numbers and Check Digit Schemes*, Washington, MAA 2001

[3] B. Williams: *Automatic Identification Systems*, Surrey (UK), Pira International Ltd 2004

[4] Več avtorjev: *For All Practical Purposes: introduction to contemporary matehematics*, New York, W. H. Freeman 2000

[5] *ISBN. Priročnik za uporabnike*, Ljubljana, NUK 2002

[6] *Uredba o načinu določanja osebne identifikacijske številke*, Uradni list št. 8/1999 (345)

Internet:

[7] http://www.gov.si/durs/index.php?lg=sl&f=05,01.html (davčna številka)

[8] http://www.ean.si

[9] http://www.ean.si/sntportal.asp?p=17 (Globalni uporabniški priročnik standardov EAN.UCC)

[10] http://www.bsi.si/html/ps/transakcijski_racun.html