České vysoké učení technické v Praze,
Fakulta dopravní

Czech Technical University in Prague,
Faculty of Transportation Science

RNDr. Štěpán Klapka, Ph.D.

Úloha detekčních kódů v železniční zabezpečovací technice

The role of Detection codes in signalling systems

Summary: This talk concentrates on the detection code utilization in signaling system. The main attention is focused on procedures and algorithms, which use the linear detection codes to reduce a risk of transportation as the part of safety control process. The talk also refers to suitable procedure for quantitative assessment of detection property of linear codes applied in signaling systems.

Souhrn: Tato přednáška je zaměřena na využití detekčních kódů v železniční zabezpečovací technice. Hlavní pozornost je věnována postupům a algoritmům, které využívají lineární detekční kódy pro omezení rizik v dopravě jako součást procesu řízení bezpečnosti. V přednášce jsou zmíněny i vhodné postupy pro kvantitativní hodnocení detekčních schopností lineárních kódů použitých v železničních zabezpečovacích zařízeních.

Klíčová slova: Detekční kódy, hodnocení bezpečnosti, železniční zabezpečovací technika.

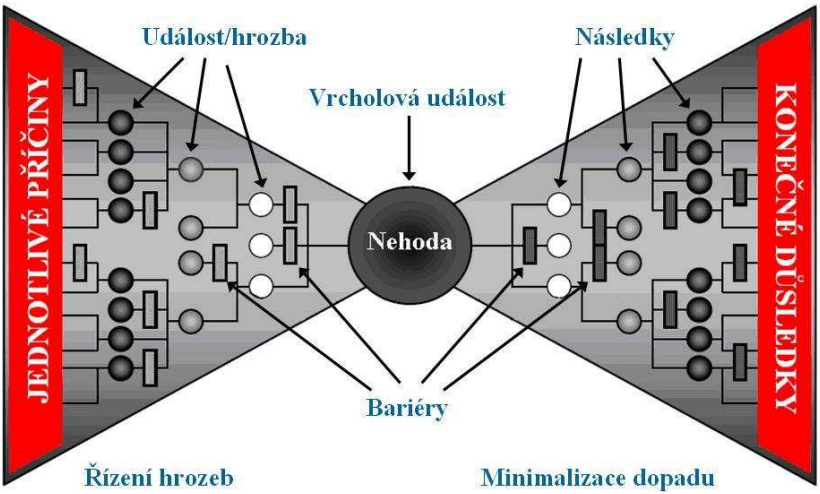
Keywords: Detection codes, safety assessment, signalling systems.

Obsah

1 Úvod

1.1 Proces řízení bezpečnosti v železniční dopravě

Smyslem procesu řízení bezpečnosti na železnici je zjednodušeně řečeno dostatečně omezit rizika spojená s železniční dopravou. Tento proces zasahuje do všech činností souvisejících s tímto způsobem dopravy a lze na něj nahlížet jako na systém bariér, které mají omezit konečné důsledky plynoucí z potenciálních hrozeb. Na následujícím schématu je patrné, že zmíněné bariéry jsou vkládány do pomyslné cesty od příčin nehod až k jejím důsledkům, a to jak z důvodu zabránění vlastní nehodové události, tak za účelem omezení následků již vzniklé nehody.



Pod abstraktní pojmem „bariéra“ si lze v uvedeném schématu představit libovolné opatření počínaje dopravními předpisy a konče technickými prostředky, jako jsou železniční zabezpečovací zařízení. Jednou z možných příčin nehodových událostí v železniční dopravě je právě taková porucha železničního zabezpečovacího zařízení, která způsobí, že toto zařízení přestane plnit svoji funkci bariéry. Touto problematikou se zabývá specifická část procesu hodnocení bezpečnosti zabezpečovacích zařízení, která se označuje jako „technická bezpečnost“. Naproti tomu „funkční bezpečnost“ se zabývá především dopravně bezpečnostními algoritmy, které jsou jinou specifickou formou výše zmíněných bariér.

Principy (algoritmy) funkční bezpečnosti velice často vychází z opatření, která vznikla jako konkrétní reakce na vzniklou nehodu. Vzhledem k této souvislosti je vývoj funkčních algoritmů nezanedbatelně ovlivněn tradicí, která se v různých částech Evropy znatelně liší.

Jedním z nástrojů, pomocí kterých se v EU postupně dosahuje určitého stupně sjednocení a harmonizace v této oblasti, jsou v první řadě direktivy evropské komise, dále evropské normy (EN 50129 [1], EN 50128 [2], EN 50159-1/2 [3][4]) a v neposlední řadě i pozvolné nasazování systému ETCS/ERTMS. Mnohé aktivity v tomto směru má v působnosti i Evropská Železniční Agentura (ERA). V současné době většina evropských norem pro oblast železnic prochází revizí v pracovních skupinách technické subkomise SC9XA CENELEC. Například ve výše jmenované subkomisi došlo v poslední době k rozhodnutí, že dvojice norem EN 50159-1/2 bude v novém návrhu sjednocena v jednu normu EN 50159, která bude platit jak pro otevřené, tak pro uzavřené přenosové systémy.

1.2 Obecná úloha detekce v železniční zabezpečovací technice

Pro objasnění úlohy detekce v železniční zabezpečovací technice je vhodné se zmínit o základních principech architektury zabezpečovacích zařízení, které vyplývají především z požadavků normy EN 50129.

Při návrhu zabezpečovacích zařízení je tedy z pohledu technické bezpečnosti nutné brát v úvahu vlivy poruchových stavů na vlastní bezpečnostní funkci zařízení. V případě uvažování vlivu ojedinělých poruchových stavů je pro systémy s vyššími požadavky na bezpečnost (SIL 3 a SIL 4) nutné zajistit, aby zůstaly bezpečné v případě jakéhokoli druhu ojedinělého náhodného poruchového stavu hardwaru, který je považován za možný. Tato zásada (princip) je známá jako **bezpečnost při poruše** (anglicky označováno jako **Fail-Safe**). Dle normy EN 50129 může být dosaženo této zásady několika různými způsoby, a to:

- inherentní (vlastní) bezpečností při poruše,
- složenou bezpečností při poruše a
- reaktivní bezpečností při poruše.

Princip inherentní (vlastní) bezpečnosti při poruše dosahuje bezpečnosti tím, že žádné hodnověrné druhy poruch jednotky (zařízení) nejsou nebezpečné. Hodnověrnost poruch musí být garantována například fyzikálními vlastnostmi použitých součástek a jejich zapojením. V tomto případě je zvládnutí poruchy (detekce a negace) zajištěno především fyzikálními zákony.

Naproti tomu složená a reaktivní bezpečnost již přímo využívá detekce pro zabránění nebezpečí (k dosažení bezpečnosti). V případě využití složené bezpečnosti je k detekci poruchových stavů použit hlasovací princip (viz norma EN 50129). V případě reaktivní bezpečnosti je rychlá a hodnověrná detekce zajištěna specializovanou jednotkou, která je k tomuto účelu navržena. Tato speciální jednotka však nevykonává přímo bezpečnostní funkci, ale jen dohlíží na správné vykonávání bezpečnostní funkce hlavní (funkční) jednotky. Současná zabezpečovací zařízení pro vysoká rizika většinou využívají všech tří principů a u některých případů se dá velice obtížně rozhodnout, o který z uvedených principů se právě jedná.

Dále se budeme zabývat obecnými vlastnostmi procesu detekce (parametry detekce), tak jak je uvažují normy EN 50129, EN 61508 [5], EN 50128 a EN 50159-1/2. Například pro techniku složené bezpečnosti při poruše je požadováno: *„(Potencionálně) Nebezpečný poruchový stav v jedné jednotce musí být detekován a negován (zvládnut) v době dostatečné k tomu, aby se zabránilo souhlasnému poruchovému stavu v druhé jednotce.“* Přesněji řečeno se požaduje, aby poruchový stav byl zvládnut dříve, než selže zvolený postup detekce (hlasování) vzhledem k další degradaci systému. V tomto požadavku jsou uvažovány pouze systémy 2oo2 a 2oo3, protože pro systém 3oo5 by druhá porucha se stejným projevem ještě nevedla k selhání hlasovacího principu.

Pro normu EN 50129 je typické, že ve svých požadavcích nezohledňuje omezenost detekčních mechanizmů a přiklání se jen ke kvalitativnímu pohledu, kdy detekce je, anebo není (viz předešlá citace). Přitom je v této normě speciálně kladen důraz na rychlost detekčního procesu (viz příloha A a D). Na druhou stranu kvantitativní pohled bere do úvahy i detekční pokrytí či stupeň nezávislosti systému vzniklých poruch. Například v normě EN 61508 se uvažuje detekční (diagnostické) pokrytí vyjadřující kvalitu detekce jako podíl na snížení pravděpodobnosti nebezpečných poruch hardwaru v důsledku provádění automatických diagnostických testů. Ve vzorcích pro vyjádření intenzity nebezpečí v části normy EN 61508-6 je dále zohledněna jak časová složka postupu detekce, tak obnova systému. Pro některé konfigurace systému (například pro 2oo3) je uvažován i faktor nezávislosti jednotek.

Nakonec ze všech norem věnovaných železniční zabezpečovací technice pouze v informativní příloze A normy EN 50159-1 lze vytušit určitou snahu o podchycení obou diskutovaných parametrů detekce (rychlosti i kvality).

Jeden z možných přímých důsledků nedokonalé detekce poruchových stavů zařízení je možnost jejich hromadění (akumulace). Některá kombinace nahromaděných poruch nakonec může vést k nebezpečnému stavu zařízení.

Systém selhávání detekčních postupů v zařízení může být velice složitý a obtížně modelovatelný proces. Vícenásobné poruchové stavy se mohou, ale zároveň nemusí, vzájemně maskovat a to jak mezi jednotkami, které pracují v systému 2oo2 nebo 2oo3, tak v jednotce samotné. Stanovení nebezpečných kombinací různých množin poruchových stavů je velice náročné a vyžaduje systematický přistup.

Na tomto místě je vhodné se zmínit i o složitosti poruchových módů u součástek s vysokou integraci, jako jsou mikroprocesory, paměti, řadiče a budiče různých typů. Možná degradace složitých polovodičových struktur klade na detekční postupy vysoké nároky. Mnozí výrobci běžných komerčních integrovaných obvodů přidávají do různých typů součástek detekční mechanizmy, které ve většině případů vycházejí z detekčních a samoopravných kódů. Detekční pokrytí těchto mechanizmů však ne vždy odpovídá současným potřebám zabezpečovacích zařízení, a je tedy nutné neporušenost (integritu) software i hardware sledovat pomocí dalších detekčních postupů, které zahrnují různé formy vestavěných testů a kontrol.

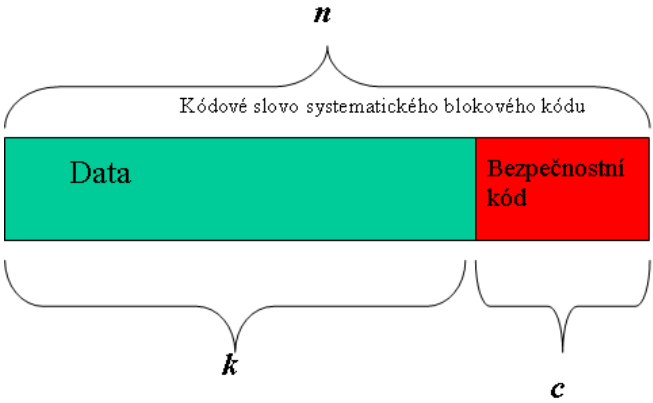
Vzhledem k tomu, že hned první porucha může postihnout právě systém kontrolních a testovacích mechanizmů, je nutné, aby návrh softwarové a hardwarové architektury byl schopen eliminovat i takové nebezpečí. Jedním z možných opatření je zajistit, aby jednotka, která není v poruše, měla možnost včas detekovat ztrátu integrity sousední jednotky. Ne vždy tento požadavek může být zajištěn pouhým hlasováním o výstupních signálech (viz systém se složenou bezpečností při poruše), ve kterých se nemusí projevit stupeň degradace jednotlivých jednotek v systému. Jedním z vhodných postupů pro ověřování integrity jednotek navzájem jsou například systémy otisků (hash funkce) rozhodujících vnitřních stavových veličin, které při hlasování zajistí větší detekční pokrytí široké škály poruchových stavů. Pokud se jako otisk dat použije lineární kód, pak je možné i analyzovat jeho detekční vlastnosti v některém z vhodných pravděpodobnostních modelů a dokonce vyčíslit pravděpodobnost kolize otisků (stejný otisk pro různá data).

2 Specifické požadavky na detekční kódy

V této části přednášky budou jen stručně shrnuty základní požadavky, které mohou být kladeny jak na detekční kódy, tak na jejich implementaci při použití v zabezpečovací technice.

Bezpečnostní požadavky na použití detekčních kódů v případě komunikací jsou stanoveny především dvojicí norem EN 50159-1/2. Obě normy, stejně jako připravovaná nová norma vzniklá jejich spojením, požadují, aby zprávy vztahující se k bezpečnosti byly chráněny bezpečnostním kódem.

Bezpečnostní kód je v těchto normách zjednodušeně definován pouze jako kontrolní část systematického blokového detekčního kódu (viz následující obrázek). V nové normě EN 50159 se však navíc připouští, že bezpečnostní kód je například kombinací lineárního detekčního kódu a konstantní části zprávy vztahující se k bezpečnosti. Tento přístup je možné zobecnit na všechny další části zprávy vztahující se k bezpečnosti, které umožní detekci ztráty integrity došlé zprávy, jako je například i sekvenční číslo zprávy.



Normativním požadavkům, které souvisí s hodnocením bezpečnostních kódů, se budeme podrobněji věnovat v následující části přednášky. Na tomto místě je vhodné se zmínit i o normativních požadavcích na detekční kódy, které souvisí s kontrolou integrity systému. Jak norma EN 50129, tak EN 50128 doporučuje věnovat pozornost oblastem samočinného testování, sledování degradace a metodám negace (zvládnutí poruchového stavu). Ve všech těchto oblastech mohou být použity detekční kódy jako základní stavební prvek a vzhledem k tomuto použití vzniká celá řada specifických požadavků. Tyto nové požadavky mohou například souviset se způsobem implementace postupů vytváření detekčních kódů (výpočet kontrolní části) a následné kontroly integrity v programovém vybavení zabezpečovacích zařízeních. V systémech, kde se dá vyloučit úmyslný útok, může být specifický datový otisk pomocí lineárního detekčního kódu znakem autenticity vytvořené zprávy. Tuto autentickou zprávu však musí umět libovolný oprávněný příjemce ověřit, ale přitom neumožnit, aby poruchou v zařízení vznikla schopnost tuto autentickou zprávu vytvořit.

3 Hodnocení detekčních kódů

Smyslem hodnocení bezpečnostních kódů je vyhovět normativním požadavkům, a to především R15 a R16 normy EN 50159-1 a článku 6.3.7.2 normy EN 50159-2. V případě připravovaného znění EN 50159 je situace obdobná. V těchto ustanovení norem se požaduje, aby bezpečnostní kód detekoval a působil jak na typické chyby vzniklé při přenosu, tak na poruchové stavy nedůvěryhodného přenosového systému. K prokázání těchto vlastností bezpečnostních kódů se využívají následující ukazatele vlastností detekčních kódů, které zde jen stručně shrnu tak, aby bylo patrné, v čem je přínos pravděpodobnostních modelů pro hodnocení detekční kvality blokových kódů.

3.1 Základní definice

Protože se dále budeme věnovat především lineárním kódům, uvedu zde jen nezbytné definice pro další výklad. Lineární kód je definován jako podprostor vektorového prostoru T^n, to znamená jako jeho podmnožina uzavřená vzhledem ke sčítání a k násobení skalárem. Kódová slova lineárního kódu jsou vektory délky *n* a součet dvou kódových slov je opět kódovým slovem. V důsledku toho každý lineární kód obsahuje nulové slovo (slovo složené ze samých nul).

Dimenze *k* lineárního kódu (jakožto dimenze lineárního podprostoru) reprezentuje počet informačních znaků kódu. Binární lineární kód délky *n* a dimenze *k*, označovaný jako lineární (*n*,*k*)-kód, má *k* informačních bitů a $c = n - k$ kontrolních (redundantních) bitů. Binární lineární (*n*,*k*)-kód má 2^k kódových slov.

Z praktického hlediska jsou významnou skupinou lineárních kódů kódy systematické. Jsou to kódy, jejichž kódové slovo vznikne prostým přidáním kontrolní části za informační slovo. Každý lineární kód je ekvivalentní s nějakým systematickým kódem, to znamená, že jej lze převést na systematický kód pouhou permutací pořadí znaků v kódových slovech.

Matice G, jejíž řádky jsou právě všechny prvky některé báze daného lineárního (*n*,*k*)-kódu C, se nazývá generující matice kódu C. Matice G má k řádků a n sloupců a plně určuje lineární kód C.

$$\left(\forall v \in T^n\right)\left((v \in C) \Leftrightarrow \left(\exists u \in T^k\right)\left(v = uG\right)\right)$$

Kontrolní matice H daného lineárního (*n*,*k*)-kódu C je matice s $n - k$ řádky a *n* sloupci, pro kterou platí: slovo délky *n* je kódovým slovem kódu C právě tehdy, když jeho součin s maticí H je nulový vektor:

$$\left(\forall v \in T^n\right)\left(v \in C \Leftrightarrow Hv^T = o^T\right)$$

Duální kód $C^{\perp}$ k danému lineárnímu (n,k)-kódu C se skládá ze všech slov délky n, která jsou ortogonální ke každému kódovému slovu kódu C:

$$\left(u \in C^{\perp}\right) \Leftrightarrow \left(\forall w \in C\right)\left(u \cdot w = 0\right)$$

Duální kód k lineárnímu (n,k)-kódu C je lineární $(n,n-k)$-kód; duálním kódem ke kódu $C^{\perp}$ je opět kód C. Generující matice G původního kódu C je kontrolní maticí duálního kódu $C^{\perp}$ a naopak kontrolní matice H původního kódu C je generující maticí duálního kódu $C^{\perp}$.

Cyklický kód je lineární kód, který je uzavřený vzhledem k cyklickému posunu: lineární kód C se nazývá cyklický, jestliže pro každé kódové slovo $(a_0, a_1, a_2, ...,a_{n-1})$ je také slovo $(a_{n-1}, a_0, a_1,...,a_{n-2})$ kódovým slovem. Cyklický posun odpovídá násobení polynomem x modulo x^n-1.

$$x \cdot \left(a_0 + a_1 x + ... + a_{n-1} x^{n-1}\right) \bmod \left(x^n - 1\right) = \left(a_0 x + a_1 x^2 + ... + a_{n-1} x^n\right) \bmod \left(x^n - 1\right) =$$

$$= a_{n-1} + a_0 x + a_1 x^2 + ... + a_{n-2} x^{n-1}$$

Polynom g(x) se nazývá generující polynom kódu C a má následující vlastnosti:

- Kód *C* se skládá právě ze všech násobků polynomu *g(x)*:
  $C = \left\{p(x)g(x); p(x) \in T \mid \left(x^n - 1\right)\right\}$
- Polynomy *g(x), xg(x), x^2g(x),...,x^kg(x)* tvoří bázi kódu *C*.
- Polynom *g(x)* dělí polynom *x^n-1* beze zbytku.

Cyklický kód je jednoznačně určen svým generujícím polynomem.

3.2 Základní ukazatele kvality detekce

Pro hodnocení efektivity (n,k) lineárního blokového kódu se používá informační poměr, který je stanoven poměrem počtu informačních znaků k a délky kódového slova n.

Nejdéle, a doposud i nejčastěji, používaným parametrem detekce souvisejícím s náhodnými chybami je minimální vzdálenost kódu. Kód s minimální vzdáleností d detekuje všechny chyby až do násobnosti *d-1*. Většina kódů však detekuje i některé chyby násobnosti vyšší. V důsledku toho se jednotlivé kódy se stejnou délkou a minimální vzdáleností mohou velmi výrazně lišit.

S minimální vzdáleností kódu rovněž souvisí definice MDS kódů (MDS - Maximum Distance Separable). Velikost minimální vzdálenosti lineárního (n,k)- kódu je totiž omezena následujícím jednoduchým vztahem (Singleton bound).

$$d \leq n - k + 1$$

Pokud nastává extrémní případ rovnosti v právě uvedeném vztahu, pak je příslušný kód MDS a má předem známé váhové rozložení.

Kromě náhodných chyb jsou další specifickou skupinou shlukové chyby. Shlukem chyb délky b se nazývá chybové slovo $e = e_1,e_2,...,e_n$, jehož všechny nenulové složky tvoří část ležící mezi b po sobě následujícími znaky.

Lineární kód objevuje shluky délky b, jestliže žádné nenulové kódové slovo není shlukem délky b. Cyklické (n,k)- kódy objevují shluky chyb délky n-k, ale neobjevují (každý) shluk chyb délky n-k+1. Jinými slovy se dá říci, že stupeň generujícího polynomu určuje maximální délku shluku, který kód vždy detekuje, a vlastní generující polynom je vzorem shlukové chyby, která nebude kódem nikdy odhalena. Obdobně jako pro náhodné chyby lze konstatovat, že většina kódů detekuje i některé shlukové chyby větší délky. V důsledku toho se jednotlivé kódy mohou velmi výrazně lišit ve své odolnosti proti shlukovým chybám i v případě, kdy mají stejnou délku a stejný stupeň generujícího polynomu.

3.3 Pravděpodobnostní modely

Protože výše uvedené ukazatele v mnoha případech nedovedou dostatečně odlišit detekční vlastnosti dvou kódů, jsou pro jejich přesnější porovnání použity pravděpodobnostní modely přenosových kanálů. Základním smyslem těchto modelů je zjednodušeným způsobem vystihnout strukturu vzorů chyb, které příslušný kód nedovede odhalit. Novým ukazatelem této vlastnosti je pravděpodobnost neodhalení chyby došlé zprávy v závislosti na pravděpodobnosti chyb v jednom znaku. Principiální zjednodušení získáme pomocí nerealistických předpokladů nezávislosti a symetrie pravděpodobnosti chyb v jednom znaku. Protože počet různých znaků se označuje písmenem Q, je výše popsaný model nazýván Q-nární Symetrický Kanál (QSC – Q-nary Symmetrical Channel).

Základní vlastnost, která dále zásadně zjednodušuje analýzu detekčních kódů, je jejich případná linearita. Pro blokové lineární detekční kódy platí, že systém vzorů neodhalitelných chyb je tvořen nenulovými kódovými slovy příslušného kódu a to nezávisle na tom, které kódové slovo je chybou postiženo. Tedy vlastní struktura nenulových kódových slov dává

nejpřesnější obraz o systému neodhalitelných modifikací a pravděpodobnost neodhalitelné chyby je pak zjednodušené měřítko struktury kódových slov a tedy i vhodnosti příslušného kódu.

Dále se budeme věnovat dvěma jednoduchým modelům, které se nejčastěji používají k teoretickému zkoumání detekčních vlastností lineárních kódů. Model Q-nárního Symetrického Kanálu (QSC– Q-nary Symmetrical Channel) je zmíněn především proto, že je obecnější formou přenosového kanálu a pro MDS kódy dává principiální výsledky z hlediska obvyklých kritérií detekce. Pravděpodobnost neodhalitelné chyby lineárního kódu v QSC je pak dána následujícím vztahem:

$$p_{ud}(\varepsilon, n, C) = \sum_{i=d}^{n} A_i^{C,n} \left(\frac{\varepsilon}{q-1} \right)^i (1-\varepsilon)^{n-i} ,$$

kde koeficienty $A_i^{C,n}$ představují počty kódových slov kódu C váhy *i* při délce kódového slova n, ε je chybovost (pravděpodobnost změny jednoho znaku) v QSC kanálu a d je minimální Hammingova vzdálenost kódu.

Nejčastěji užívaný je však Binární Symetrický Kanál (BSC – Binary Symmetrical Channel), který je jako přijatelný model pro prokazování detekčních vlastností doporučen i v připravovaném znění normy EN 50159. Pravděpodobnost neodhalitelné chyby lineárního kódu v BSC je dána následujícím vztahem:

$$p_{ud}(p_e, n, C) = \sum_{i=d}^{n} \left(A_i^{C,n} p_e^i (1-p_e)^{n-i} \right),$$

kde opět koeficienty $A_i^{C,n}$ představují počet kódových slov váhy *i* při délce kódového slova *n*, p_e je bitová chybovost (pravděpodobnost změny jednoho bitu) v BSC kanálu a d je minimální Hammingova vzdálenost kódu.

3.4 Výpočet váhového rozložení kódu

Z výše uvedených rovnic vyplývá, že pro výpočet pravděpodobnosti neodhalení chyby v modelu BSC je nutné znát váhové rozložení kódu. Přirozeným postupem, jak váhové spektrum spočítat, je vygenerování všech nenulových kódových slov a vypočítání jejich vah. Pokud uvažujeme binární lineární (*n*,*k*)-kód, pak počet všech kódových slov je 2^k a pro k>48 je takový výpočet obtížně uskutečnitelný.

Z těchto důvodů se váhové spektrum počítá nepřímo pomocí váhového spektra duálního kódu, jehož výpočet přímou metodou je závislý na

velikosti redundance *n-k* (2^{n-k} duálních kódových slov). Vztah, který provazuje váhové spektrum primárního a duálního kódu pomocí jejich váhových polynomů, se nazývá identita McWilliamsové. Následující rovnice je jeden z možných tvarů identity pro binární kód.

$$2^k \sum_{i=0}^{n} \left(B_i^{C^\perp,n} x^i \right) = (1+x)^n \sum_{i=0}^{n} \left(A_i^{C,n} \left(\frac{1-x}{1+x} \right)^i \right),$$

kde koeficienty $A_i^{C,n}$ a $B_i^{C^\perp,n}$ představují počet kódových slov váhy *i* primárního a duálního kódu při délce kódového slova *n*. Pomocí transformací (substitucí) a úprav výše uvedené rovnice lze dosáhnout toho, že lze pomocí váhového rozložení duálního kódu vyjádřit pravděpodobnost neodhalitelné chyby v BSC:

$$p_{ud}(p_e, n, C) = 2^{k-n} \sum_{i=0}^{n} \left(B_i^{C^\perp,n} (1-2p_e)^i \right) - (1-p_e)^n$$

Pro urychlení výpočtu kódových slov cyklického (případně zkráceného) duálního kódu lze použít metodu, se kterou poprvé přišli Fujiwara et al. v článku [6] a která byla později rozšířena Castagnolim et al. v článku [7]. Tato metoda umožňuje efektivně vypočítat váhové rozložení při obdobné náročnosti i pro relativně veliká n.

3.5 Kritéria pro hodnocení detekčních kódů

V následujících odstavcích jsou uvedeny definice „správného“ (proper) a „dobrého“ (good) kódu, tak jak jsou užívány v odborné literatuře v případě modelu BSC. Obdobným postupem lze definovat obdobné pojmy pro model QSC.

Definice 3.1

Řekneme, že binární kód je „správný“, pokud pravděpodobnost neodhalitelné (nedetekované) chyby $p_{ud}(p_e)$ v modelu BSC pro hodnoty p_e menší než jedna polovina je neklesající (monotonní) funkcí pravděpodobnosti bitové chyby p_e.

Definice 3.2

Řekneme, že binární kód s c kontrolními bity je „dobrý“, pokud pro hodnoty pravděpodobnosti bitové chyby p_e v modelu BSC menší než jedna polovina pravděpodobnost neodhalitelné (nedetekované) chyby $p_{ud}(p_e)$ je menší než 2^{-c}. (Některé zdroje uvádějí, například v článku [8], že pro „dobrý“ kód je stanovena mez hodnotou $p_{ud}(1/2)=2^{-n}(2^k-1)$)

V obecném případě (nebinárním), při použití modelu QSC, se dále bude používat pojem q-správný nebo q-dobrý kód. Správnost kódu pro konstrukční délku nezaručuje, že kód bude správný v případě jeho zkrácení. Obecně lze konstatovat, že monotonie funkce $p_{ud}(p_e)$ jakéhokoliv lineárního kódu je zaručena až do hodnoty $p_e=d/n$ (relativní minimální vzdálenost kódu). Pro ilustraci jsou v následující části přednášky zmíněny některé postupy pro ověření „dobrost“ či „správnost“ kódu.

Na tomto místě je vhodné poznamenat, že pro kalkulace, prokazující, že bude dosaženo kvantitativních bezpečnostních cílů, nejsou právě zmíněná kritéria nezbytně potřebná. I v připravované verzi normy EN 50159 je totiž požadováno, aby pro zmíněné kalkulace bylo použito nejhorších hodnot pro pravděpodobnost neodhalitelné chyby v příslušném modelu přenosového kanálu. Tedy je možné použít i „nedobrý“ kód, ale je nutné pro něj spočítat příslušné maximální hodnoty analýzou extrémů funkce $p_{ud}(p_e)$.

3.6 Postupy pro ověření kritérií

V některých případech (pro malé d) je možné použít následující jednoduchý spodní odhad. Buď C lineární (n,k)-kód s minimální vzdáleností $d<n/2$. Jestliže je známa hodnota $A_d^{C,n}$ (nebo alespoň její dolní odhad), můžeme porovnat maximální hodnotu výrazu $A_d^{C,n} p_e^d (1-p_e)^{n-d}$ (která je dosažena pro $p_e= d/n$) s nejvyšší přijatelnou hodnotou p_{ud}.

Právě popsaným způsobem lze tedy vyloučit to, že hodnocený kód není dobrý, a proto nemůže být ani správný. Pokud jsou dostupné odhady dalších koeficientů váhového spektra, lze odhad zpřesnit, ale správnost kódu nelze tímto způsobem prokázat.

Známe-li minimální vzdálenost primárního kódu a jeho duálního kódu, můžeme v některých případech využít výsledků uvedených v článku [8]. V dalším textu bude symbolem $\lfloor x \rfloor$ označována dolní celá část reálného čísla x (“zaokrouhlení dolů”), symbolem $\lceil x \rceil$ jeho horní celá část (“zaokrouhlení nahoru”).

Tvrzení 3.1

Buď C lineární (n,k)-kód s minimální vzdáleností d a nechť $d^\perp$ je minimální vzdálenost duálního kódu $C^\perp$.

a) Jestliže platí $d^\perp \geq \lfloor n/2 \rfloor + 1$, pak kód C i kód $C^\perp$ jsou „správné“.

b) Analogicky jsou oba kódy „správné“ i v případě, že je splněna nerovnost $d \geq \lfloor n/2 \rfloor + 1$.

c) Jestliže platí $\lceil n/3 \rceil + 1 \leq d^{\perp} \leq \lfloor n/2 \rfloor$, pak je kód C „správný“ na intervalu $\left[\frac{n+1-2d^{\perp}}{n-d^{\perp}}, \frac{1}{2} \right]$.

d) Jestliže vedle $\lceil n/3 \rceil + 1 \leq d^{\perp} \leq \lfloor n/2 \rfloor$ platí $n\left(n+1-2d^{\perp}\right) \leq \left(n-d^{\perp}\right)d$, je kód C „správný“.

Protože výše uvedené tvrzení lze použít pro ověření správnosti na intervalu [0,1/2] jen ojediněle, je dále zmíněn postup ověřování správnosti kódu pomocí rozšířených binomických momentů kódu (podrobněji viz [9] a [10]).

Nechť C je q-nární lineární (*n*,*k*)-kód s váhovým vektorem-rozložením (A_0,..., A_n). Rozšířený binomický moment kódu A* je definován následujícím předpisem:

$$A_0^* = 0,\ A_l^* = \sum_{i=0}^{l} \frac{l(l-1)\cdots(l-i+1)}{n(n-1)\cdots(n-i+1)} A_i, \quad l = 1..n.$$

Rozšířené binomické momenty A* kódu C a rozšířené binomické momenty B* duálního kódu $C^{\perp}$ jsou svázány následujícími vzorci:

$$A_l^* + 1 = q^{l-n+k}\left(B_{n-l}^* + 1\right), \quad l = 0..n.$$

$$B_l^* + 1 = q^{l-k}\left(A_{n-l}^* + 1\right), \quad l = 0..n.$$

Právě uvedené vzorce jsou obdobou již uvedené identity McWilliamsové. Tedy pomocí duálních rozšířených binomických momentů vypočítáme rozšířené binomické momenty primárního kódu. Pomocí následujících tvrzení pak můžeme rozhodnout o vlastnostech kódu.

Tvrzení 3.2

Buď C q-nární lineární (*n*,*k*)-kód s minimální vzdáleností *d*, minimální vzdálenost $d^{\perp}$ duálního kódu $C^{\perp}$ a platí $d + d^{\perp} \leq n$. Jestliže dále platí $A_l^* \geq qA_{l-1}^*$ pro $l = d+1,..., n-d^{\perp}+1$, pak je kód C q-správný.

Počet nerovností, které je nutno zkoumat, je roven $n-\left(d+d^{\perp}\right)+1$ a je tedy tím menší, čím je součet $d+d^{\perp}$ vyšší. Na tomto místě je vhodné zmínit, že pro MDS kódy, které jsou q-správné platí, že $d+d^{\perp}=n+2$. Obdobná věta platí i pro rozšířené binomické momenty duálního kódu $C^{\perp}$.

Tvrzení 3.3

Jestliže platí $B_{n-l}^{*} \geq B_{n-l+1}^{*} - q^{n-k-l}\left(q-1\right)$ pro $l=d+1,...,n-d^{\perp}+1$, pak je kód C q-správný.

Není tedy nezbytné pro ověření správnosti počítat rozšířené binomické momenty původního kódu.

Slabší podmínka platí pro „dobrý“ kód:

Tvrzení 3.4

Nechť C je q-nární lineární (*n*,*k*)-kód s minimální vzdáleností d, minimální vzdálenost $d^{\perp}$ duálního kódu $C^{\perp}$ a platí $d+d^{\perp} \leq n$, rozšířené binomické momenty kódu C jsou A_l^{*} a rozšířené binomické momenty duálního kódu $C^{\perp}$ jsou B_l^{*}. Jestliže platí $A_l^{*} q^{-l} \leq q^{-n}\left(q^{k}-1\right)$ pro $l=d,...,n-d^{\perp}$, nebo $q^{-n+l} B_{n-l}^{*} \leq q^{-k} - q^{-n-k+1}$ pro $l=d,...,n-d^{\perp}$, pak je kód C q-dobrý.

Při použití právě popsaných postupů ověřování vlastností kódů se ukazuje několik dalších úskalí, a to zaprvé, že složitost výpočtu (ověření všech nerovností) má kvadratickou závislost na délce kódového slova n a zadruhé, s rostoucím n se zvyšují požadavky na přesnost výpočtu, což v důsledku opět o řád zvyšuje složitost výpočtu. Tedy pro kódy s malým n (pro obvyklou velikost přenášených zpráv) je to použitelný postup ověření jejich vlastností, který s rostoucím n ztrácí smysluplné použití (například pro kódy kontroly integrity pamětí).

Určitého zefektivnění výpočtu lze dosáhnout v případě použití kritéria dle výše uvedeného Tvrzení 3.3 tím, že se využije monotonie posloupnosti rozdílů ($B_{i+1}^{*} - B_i^{*} \geq B_i^{*} - B_{i-1}^{*}$) a při výpočtu v pohyblivé řádové čárce se vyhodnocuje spodní a horní odhad hodnoty rozdílu rozšířených binomických momentů vzhledem k vlivu zaokrouhlovacích chyb na výpočet. Tímto způsobem je dosažitelná analýza detekčních vlastností

zkráceného cyklického kódu s redundancí 32 bitů až do velikosti dat v řádech MB při použití dvojnásobné přesnosti.

4 Závěr

V této přednášce byla hlavní pozornost věnována postupům a algoritmům, které využívají lineární detekční kódy pro omezení rizik v dopravě jako součást procesu řízení bezpečnosti. V přednášce byly zmíněny některé vhodné postupy pro kvantitativní hodnocení detekčních schopností lineárních kódů použitých v železničních zabezpečovacích zařízeních. Některé výsledky získané popsanými postupy byly již prezentovány například viz [11] a [12]. Vzhledem k možnostem současné výpočetní techniky se dá očekávat, že uvedené postupy se v blízké budoucnosti stanou nezbytnou součástí kalkulací pro prokázání dosažení kvantitativních cílů bezpečnosti železničních zabezpečovacích zařízení.

Reference

[1]	ČSN EN 50129 - Drážní zařízení – Sdělovací a zabezpečovací systémy a systémy zpracování dat – Elektronické zabezpečovací systémy
[2]	ČSN EN 50128 Sdělovací a zabezpečovací systémy a systémy zpracování dat – Software pro drážní řídící a ochranné systémy.
[3]	ČSN EN 50159-1 Sdělovací a zabezpečovací systémy a systémy zpracování dat – Bezpečná komunikace v uzavřených systémech.
[4]	ČSN EN 50159-2 Sdělovací a zabezpečovací systémy a systémy zpracování dat – Bezpečná komunikace v otevřených systémech.
[5]	ČSN EN 61508 Funkční bezpečnost elektrických/elektronických/ programovatelných elektronických systémů souvisejících s bezpečností
[6]	Fujiwara T., Kasami T., Kitai A. and Lin S.: On the Undetected Error Probability for Shortened Hamming Codes, IEEE Transactions on Communications, Vol. 33, pp. 570-574, June 1985.
[7]	Castagnoli G., Brauer S., Hermann M.: Optimization of Cyclic Redundancy-Check Codes with 24 and 32 Parity Bits. IEEE Transactions on Communications, Vol. 41, pp. 883-892, June 1993.

[8]	Dodunekova R. and Nikolova E.: Sufficient Conditions for Monotonicity of the Undetected Error Probability for Large Channel Error Probabilities, Problems of Information Transmission, Vol. 41, No. 3, 2005, pp. 187-198. Translated from Problemy Peredachi Informatsii, No. 3, 2005, pp. 3-16. Original Russian Text Copyright © 2005 by Dodunekova, Nikolova.
[9]	Dodunekova R.: On the Binomial Moments of Linear Codes end Undetected Error Probability. Preprint No. 2002:49, Department of Mathematics, Chalmers University of Technology and Göteborg University, 2002.
[10]	Dodunekova, R.: Extended Binomial Moments of a Linear Code and the Undetected Error Probability, Probl. Peredachi Inf., 2003, vol. 39, no. 3, pp. 28-39 [Probl. Inf. Trans. (Engl. Transl.), 2003, vol. 39, no. 3, pp. 255-265].
[11]	Harlenderová M., Kárná L., Klapka Š.: Calculation of detection properties in a binary symmetrical channel. In: Formal Methods for Automation and Safety in Railway and Automotive Systems (Forms/Format 2007 proceedings).
[12]	Kárná, L., Klapka Š., Harlenderová M.: Quantitative Assessment of Safety Code. Proc. FORMS/FORMAT 2008, l'Harmattan, Budapest, pp 249-255, 2008.

RNDr. Štěpán Klapka, Ph.D.

Výzkumný pracovník specialista (AŽD Praha s.r.o), narozen 15.11.1963

Vzdělání:

1982 – 1987 MFF UK obor - přibližné a numerické metody (diplomová práce na téma – Spolehlivost odhadu lokální chyby)

1995 – 2002 Doktorandské studium MFF UK – vědecko-technické výpočty (disertační práce na téma – Markovovské modelování v zabezpečovací technice)

Zaměstnání:

1987 – 1993 Výzkumný ústav matematických strojů (náplň práce – matematické modelování v následujících oblastech:

- polovodičové struktury, identifikace parametrů pro dynamické obvodové modely,

- rozeznávání řeči (spektrální analýza, lineární prediktivní kódování),

- náhodné generátory, metoda Monte Carlo,

- vedení tepla.

1990 – 1995 poloviční úvazek na KNM MFF UK (náplň práce – seminář metody konečných prvků PLTMG, studentské projekty ROBOT I-III)

1993 – 1995 Dita s.r.o (náplň práce – analytik programátor, návrh pseudonáhodných generátorů pro výherní automaty)

1995 Czech Alarms s.r.o (náplň práce – analytik programátor, komunikační protokoly EZS)

1996 – Výzkum a vývoj AŽD Praha s.r.o, náplň práce – matematické modelování, návrh programového vybavení v mnoha oblastech:

- návrh komunikačních protokolů pro kritické aplikace (železniční zabezpečovací systémy), návrh SW centrální jednotky ABE-1,

- teorie kódování – návrh detekčních kódů pro kritické aplikace (ESA11, ABE-1, LS06, STM, IRI),

- kvantitativní hodnocení kvality detekce na modelu BSC/QSC,

- návrh a výpočet stochastických modelů (CTMC/DTMC) pro hodnocení bezpečnosti (Markovovské řetězce s diskrétním a spojitým časem),

Pedagogické aktivity:

Podíl na přednáškách a cvičení na FD ČVUT: Železniční zabezpečovací systémy, Matematické algoritmy, Teorie hromadné obsluhy

Vedení doktorandů (3)

Obhájených diplomových prací (3)

Člen Oborové rady PGS (Inženýrská informatika v dopravě a spojích)

Jiné aktivity a ocenění:

Zlatá medaile na mezinárodním veletrhu v Brně (2001) za elektronický autoblok ABE-1, AŽD Praha s.r.o.

GRANT:

- COPERNICUS 94-95 - Control of Robot motion

- MD ČR - PROJEKT 1F43D/019/030 – „Bezpečnostní politika pro datové přenosy zabezpečovacích zařízení v železniční dopravě“ (2004-2007)