

**SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY**

**DEŠIFROVANIE V QC-LDPC MCELIECEOVOM
KRYPTOSYSTÉME
TÍMOVÝ PROJEKT**

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

DEŠIFROVANIE V QC-LDPC MCELIECEOVOM
KRYPTOSYSTÉME
TÍMOVÝ PROJEKT

Študijný program: Aplikovaná informatika
Číslo študijného odboru: 2511
Názov študijného odboru: 9.2.9 Aplikovaná informatika
Školiace pracovisko: Ústav informatiky a matematiky
Vedúci záverečnej práce: Ing. Viliam Hromada, PhD., Mgr. Tomáš Fabšič, PhD.

Bratislava 2019

Forks

Podakovanie

Chceli by sme sa podakovať vedúcim tímového projektu Mgr. Tomášovi Fabšičovi, PhD. a Ing. Viliamovi Hromadovi, PhD. za odborné vedenie počas práce na tímovom projekte, ako aj za cenné rady ohľadom implementácie riešenia.

Obsah

Úvod	1
1 Ponuka	2
1.1 Riešiteľský kolektív	2
1.2 Anotácia tímového projektu	4
1.3 Motivácia	5
1.4 Organizácia projektu	5
2 Teoretická časť	6
2.1 Úvod do problematiky	6
2.2 Lineárne kódy	6
2.3 McEliece kryptosystém	9
2.3.1 Generovanie kľúčov	9
2.3.2 Šifrovanie	9
2.3.3 Dešifrovanie	9
2.4 QC-LDPC McEliece	11
2.5 Bit-flippingové algoritmy	13
3 Praktická časť	15
Záver	20
Zoznam použitej literatúry	21
Prílohy	I
A Štruktúra elektronického nosiča	II
B Používateľská príručka	III

Zoznam obrázkov a tabuliek

Obrázok 1	Schéma McEliece kryptosystému [5]	10
Obrázok 2	Tannerov graf na zistenie najpodozrivejšieho bitu	13
Obrázok 3	UML diagram - dešifrovanie v QC-LDPC	15
Obrázok 4	UML diagram - dešifrovanie v QC-MDPC	16
Obrázok 5	Bit-flippingový algoritmus A3 [7]	17
Obrázok 6	Bit-flippingový algoritmus B2 [7]	18
Obrázok 7	Tabuľka výsledkov experimentu	19
Obrázok 8	Graf chybovosti dešifrovania	19

Zoznam skratiek a značiek

GF - Galois Field

RSA - Rivest–Shamir–Adleman

QC - Kvázi-cyklická (matica)

LDPC - Low density parity check

MDPC - Moderate density parity check

UPC - Unsatisfied Parity-check Equation

Úvod

Na predmete Tímový projekt sme si vybrali tému, ktorá v poslednom čase rezonuje nielen vo vedeckej komunite, a to *post-quantová* kryptografia. Hoci existuje viacero skupín kryptosystémov odolných voči kvantovému počítaču, my sme sa zamerali na tie, ktoré sú založené na tzv. *dekódovacím probléme*. Naša práca nesie názov *Dešifrovanie v QC-LDPC McElieceovom kryptosystéme* a hlavnou úlohou bolo experimentálne porovnanie dvoch spôsobom dešifrovania v tomto kryptosystéme.

Riešiteľský kolektív pozostával z piatich študentov aplikovanej informatiky pod vedením Mgr. Tomáša Fabšiča, PhD. a Ing. Viliama Hromadu, PhD. Pretože problematika, na ktorú sa zameriaval tento tímový projekt bola pre nás nová, prvou úlohou bolo teda naštudovanie si princípov fungovania QC-LDPC McElieceovho kryptosystému. Po získaní prvotných poznatkov z oblasti sme si rozdelili úlohy na projekte. Každý mal za úlohu spracovať inú časť teórie. Ďalej sme sa museli oboznámiť s implementáciou QC-LDPC McElieceovho kryptosystému v knižnici *BitPunch* a doplniť ju o ďalšiu funkcionálnosť. Pri procese tvorby praktickej časti sme sa museli s vedúcimi projektu niekoľkokrát stretnúť a konzultovať ďalší postup.

Jednou z úloh tímového projektu bolo aj vypracovanie dokumentácie, ktorá objasňuje riešenie zadania. Práca sa skladá z dvoch hlavných častí. V teoretickej časti čitateľovi predkladáme nevyhnutné matematické základy na pochopenie problematiky projektu. Následne vysvetľujeme konštrukciu McElieceovho kryptosystému. Praktická časť obsahuje popis implementácie, ako aj zistené výsledky z experimentálnej časti.

1 Ponuka

Riešiteľský tím v zložení: Bc. Nikoleta Furičková, Bc. Juraj Karásek, Bc. Matej Ohradzanský, Bc. Peter Radvan a Bc. Lukáš Štrba na tomto mieste predkladá záväznú ponuku na riešenie problému s pracovným názvom: *Dešifrovanie v QC-LDPC McElieceovom kryptosystéme*. Projekt budeme riešiť pod vedením Mgr. Tomáša Fabšiča, PhD. a Ing. Viliama Hromadu, PhD. v rámci predmetu **Tímový projekt**.

V časti **Riešiteľský kolektív** stručne predstavíme členov tímu. Zameriame sa na ich schopnosti a skúsenosti, ktoré súvisia s problematikou tímového projektu. Zadanie a úlohy, ktoré z neho plynú budú opísané v časti **Anotácia tímového projektu**. Nasledovať bude časť, v ktorej vyjadríme našu dôveru v post-quantovú kryptografiu a jej využitia v *blízkej* budúcnosti. Ponuku ukončíme popisom organizácie projektu.

1.1 **Riešiteľský kolektív**

Tím pozostáva z piatich študentov aplikovanej informatiky na Fakulte elektroniky a informatiky Slovenskej Technickej Univerzity. Členovia tímu sú spolužiaci už 4 roky a panujú medzi nimi priateľské vzťahy. Pracovný názov nášho kolektívu je *Forks*.

Bc. Nikoleta Furičková

Pozícia v tíme: Analytička

Je absolventkou bakalárskeho štúdia na Fakulte elektrotechniky a informatiky Slovenskej Technickej Univerzity v Bratislave, v študijnom programe Aplikovaná informatika, odbor Bezpečnosť informačných systémov. Bakalárske štúdium ukončila vypracovaním bakalárskej práce s názvom: *Invertovateľnosť blokovo cyklických matíc*. Bakalársku prácu robila pod vedením Mgr. Tomáša Fabšiča, PhD., pričom získala cenné skúsenosti a poznatky, ktoré sa dajú dobre využiť pri riešení tímového projektu.

Bc. Juraj Karásek

Pozícia v tíme: Developer

Je absolventom bakalárskeho štúdia na Fakulte elektrotechniky a informatiky Slovenskej Technickej Univerzity v Bratislave, v študijnom programe Aplikovaná informatika, odbor

Bezpečnosť informačných systémov. Bakalárske štúdium ukončil vypracovaním bakalárskej práce s názvom: *Porovnanie vybraných mechanizmov distribuovaného konsenzu z hľadiska IT bezpečnosti*.

Bc. Matej Ohradzanský

Pozícia v tíme: Developer

Je absolventom bakalárskeho štúdia na Fakulte elektrotechniky a informatiky Slovenskej Technickej Univerzity v Bratislave, v študijnom programe Aplikovaná informatika, odbor Bezpečnosť informačných systémov. Bakalárske štúdium ukončil vypracovaním bakalárskej práce s názvom: *Lúštenie substitučných šifrier*. Bakalársku prácu robil pod vedením prof. Ing. Pavla Zajaca, PhD. a v apríli tohto roku sa zúčastnil SVOČ-ky (*Študentská vedecká a odborná činnosť*), pričom sa umiestnil na prvom mieste vo svojej kategórii.

Bc. Peter Radvan

Pozícia v tíme: Web Developer

Je absolventom bakalárskeho štúdia na Fakulte elektrotechniky a informatiky Slovenskej Technickej Univerzity v Bratislave, v študijnom programe Aplikovaná informatika, odbor Bezpečnosť informačných systémov. Bakalárske štúdium ukončil s vyznamenaním. Pod vedením Ing. Viliama Hromadu, PhD. vypracoval bakalársku prácu s názvom: *ABC kryptosystém*, pri tvorbe ktorej získal množstvo vedomostí z oblasti post-quantovej kryptografie, ktoré predstavujú značný náskok v analytickej časti projektu.

Bc. Lukáš Štrba

Pozícia v tíme: Analytik a vedúci riešiteľského kolektívu?

Je absolventom bakalárskeho štúdia na Fakulte elektrotechniky a informatiky Slovenskej Technickej Univerzity v Bratislave, v študijnom programe Aplikovaná informatika, odbor Bezpečnosť informačných systémov. Bakalárske štúdium ukončil s vyznamenaním. Pod vedením Ing. Viliama Hromadu, PhD. vypracoval bakalársku prácu s názvom: *Kubický ABC kryptosystém*, pri tvorbe ktorej získal množstvo vedomostí z oblasti post-quantovej kryptografie. Tieto poznatky sú stále aktuálne a predstavujú dobrý základ potrebný na úspešné vypracovanie projektu.

1.2 Anotácia tímového projektu

Pokrok vo vývoji kvantového počítača má vážne dôsledky aj pre kryptografiu. Je známe, že dostatočne výkonné kvantové počítače budú vedieť efektívne riešiť problém faktorizácie čísla na prvočísla a problém diskrétného logaritmu. Na náročnosti riešenia týchto problémov je založená bezpečnosť v súčasnosti používaných asymetrických kryptosystémov (napríklad RSA). To znamená, že v prípade existencie dostatočne výkonného kvantového počítača by súčasné asymetrické kryptosystémy už neboli bezpečné. Niektoré odhady hovoria, že takto výkonné kvantové počítače by mohli existovať už o 10 rokov. Je preto dôležité, pracovať na vývoji nových asymetrických kryptosystémov, ktoré budú odolné voči útokom kvantového počítača, a ktoré by mohli nahradiť súčasné asymetrické kryptosystémy. Na dôležitosť tejto témy upozornil aj americký inštitút pre štandardy a technológiu *NIST* v správe Report on Post-Quantum Cryptography. *NIST* zároveň vyhlásil súťaž Post-Quantum Cryptography Standardization Process s cieľom navrhnúť nové kryptografické štandardy odolné voči kvantovým počítačom. Do súťaže prišlo vyše 60 návrhov kryptosystémov. Tieto návrhy budú v najbližších rokoch verejne analyzované vedeckou komunitou s cieľom vybrať najlepších kandidátov.

Viacero návrhov zaslaných do súťaže je založených na QC-LDPC McElieceovom kryptosystéme. Jedná sa o variant McElieceovho kryptosystému, v ktorom sa využívajú QC-LDPC kódy. Pri šifrovaní sa správa prevedie na kódové slovo QC-LDPC kódu a ku kódovému slovu sa pridá chyba. Pri dešifrovaní je potrebné túto chybu odstrániť použitím tajného kľúča, ktorý obsahuje matice H a Q . Na odstránenie chyby je možné použiť dve metódy: buď sa použije dekódovací algoritmus pre QC-LDPC kódy, ktorý využíva maticu H , alebo sa použije dekódovací algoritmus pre QC-MDPC kódy, ktorý používa maticu H^*Q . Cieľom práce je porovnať efektívnosť týchto dvoch metód. Úlohy, ktoré z toho vyplývajú sú nasledovné:

- naštudovanie si princípov fungovania QC-LDPC McElieceovho kryptosystému
- oboznámenie sa s implementáciou QC-LDPC McElieceovho kryptosystému v knižnici BitPunch (implementácia je v jazyku C)
- porovnanie efektívnosti dvoch vyššie spomenutých metód na odstránenie chyby zo správy zašifrovanej QC-LDPC McElieceovým kryptosystémom

1.3 Motivácia

Pre projekt sme sa rozhodli preto, že si uvedomujeme hrozbu, ktorú by predstavovalo zostrojenie dostatočne silného kvantového počítača pre dnešný svet. Dôkazom aktuálnosti tejto problematiky je aj fakt, že existujú viaceré medzinárodné projekty organizované napr. *NATO* alebo *VEGA*, do ktorých je zapojený aj Ústav informatiky a matematiky FEI STU. Skúmanie rôznych post-quantových kryptosystémov je nesmierne dôležité a sme radi, že sa toho môžeme zúčastniť. Získané vedomosti môžu predstavovať cenný zdroj nových informácií pre členov tímu, ktorí sa už podrobnejšie stretli s post-quantovou kryptografiou a zároveň dobrý základ pre členov, ktorí sú noví a vedomosti iba zbierajú. Veríme, že naše výsledky z tímového projektu pomôžu k napredovaniu výskumu v tejto oblasti.

1.4 Organizácia projektu

Z diskusie členov riešiteľského kolektívu so zadávateľmi tímového projektu vyplynulo, že jediný možný termín pravidelných konzultácií je utorok ráno (cca 08:00). S návrhom súhlasili všetci prítomní. Neprítomný Bc. Peter Radvan s časom stretnutia dodatočne súhlasil. Z pravidelných stretnutí sa priebežne tvoria zápisnice, ktoré sú spolu s ostatnými materiálmi a informáciami dostupné na stránke projektu. Súčasťou vypracovania projektu je aj odovzdanie dokumentácie, v ktorej bude opísaný matematický model QC-LDPC McElieceovho kryptosystému, ako aj nami zistené výsledky meraní. Na vypracovanie analytickej časti projektu neuvažujeme nad žiadnymi špeciálnymi hardvérovými požiadavkami, ale v prípade spúšťania väčšieho množstva výpočtov uvažujeme nad využitím klastra STU v Bratislave.

2 Teoretická časť

2.1 Úvod do problematiky

Pri prenose informácií vznikajú prirodzene rôzne otázky. Nezachytáva niekto našu komunikáciu, nemení nám niekto obsah našich správ? Ako odpoveď na tieto otázky sa dáta začali rôzne modifikovať do podoby, ktorej by tretia strana nebola schopná rozumieť t.j. *šifrovať*. Pomocou šifrovania vieme informáciu bezpečne dostať od odosielateľa k prijímateľovi. Kryptografia ako vedná disciplína sa časom vyvíjala. Postupne vznikali nové a bezpečnejšie kryptosystémy. Avšak zostrojením dostatočne silného kvantového počítača by bolo možné tieto kryptosystémy prelomiť. Súčasná asymetrická kryptografia sa spolieha na problémy z teórie čísel, ako sú napríklad *faktorizácia prvočísel* a výpočet *diskrétného logaritmu*. Tieto problémy by boli podľa P. W. Shora [11] riešiteľné v polynomiálnom čase na kvantovom počítači. Tým sa stávajú asymetrické kryptosystémy a podpisové schémy, ako napr. RSA a DSA zraniteľné a prelomiteľné [3].

Našťastie, existuje skupina matematických problémov, označované ako *NP-úplné* problémy, pre ktoré nepoznáme *efektívny* algoritmus, ktorý by ich dokázal riešiť na klasickom ani na kvantovom počítači. Je teda rozumné skúmať takéto problémy a na nich zakladať nové kryptosystémy odolné voči útokom na kvantovom počítači. Veda, ktorá sa zaoberá touto problematikou sa nazýva *post-quantová kryptografia*. Jedným z takýchto problémov je aj tzv. *dekódovací problém*, ktorý sa zaoberá dekódovaním náhodného lineárneho kódu.

2.2 Lineárne kódy

Pri prenose správy prostredníctvom komunikačného kanálu môže nastať chyba prenosu a preto príjemnca nedostane pôvodnú odoslanú správu. Riešením tohto problému sa zaoberá *teória kódovania*. Pridaním *redundantnej* informácie je možné *detekovať*, alebo dokonca *opraviť* určitý počet chýb. Takto sa správa dĺžky k rozšíri o r redundantných bitov a celková dĺžka správy narastie na $k + r$. Ďalej budú vysvetlené základné pojmy a definície.

Lineárny kód je k rozmerný lineárny podpriestor v F_q^n . Označujeme ho ako $[n, k]$. Každé kódové slovo lineárneho kódu C má dĺžku n a samotný lineárny kód C má dimenziu k . Konečné pole F_q obsahuje q prvkov. Počet kódových slov je q^k [2].

Lineárny kód vieme reprezentovať prostredníctvom *generujúcej* alebo *kontrolnej* matice. Generujúca matica G má veľkosť $k \times n$. Riadky matice sú tvorené bázou pre lineárny kód C . Báza obsahuje k lineárne nezávislých vektorov z vektorového priestoru F_q^n . Lineárny kód vieme vygenerovať nasledujúcim predpisom:

$$C = \{aG | a \in F_q^k\} \quad (1)$$

Štandardná forma generujúcej matice G má tvar $G = [I_k | P]$, kde I_k reprezentuje jednotkovú maticu veľkosti $k \times k$ a matica P má rozmery $k \times r$.

Vieme, že existuje $n - k$ lineárne nezávislých vektorov, ktoré sú kolmé na všetky vektory z lineárneho kódu C a teda lineárny kód vieme zapísať ako:

$$C = \{v \in F_q^n : Hv^T = \bar{0}\} \quad (2)$$

Maticu H , ktorá spĺňa (2) nazývame *kontrolná matica*. Pre lineárny kód C existuje viacero kontrolných matíc.

Uvedme príklad generujúcej a kontrolnej matice binárneho lineárneho kódu $[7,4]$. Generujúca matica G lineárneho kódu je definovaná štyrmi bázovými vektormi, ktoré sa nachádzajú v 7-rozmernom binárnom priestore. Počet všetkých možných kódových slov je 2^4 .

$$\text{Generujúca matica } G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

$$\text{Kontrolná matica } H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Lineárny kód máme definovaný a môžeme vytvoriť kódové slovo. Napríklad, nech vektor $v = (1, 0, 1, 0)$. Takto vytvorený vektor vynásobíme maticou G a dostaneme kódové slovo $k = (1, 0, 1, 0, 1, 0, 1)$ lineárneho kódu C . Správnosť vytvorenia kódového slova k si vieme overiť práve pomocou kontrolnej matice H tak, že po vynásobení dostaneme nulový vektor.

$$H \cdot k^T = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Nižšie uvádzame ďalšie potrebné informácie dôležité pre problematiku tímového projektu. Kontrolné matice H pre binárny kód, môžu mať rôznu hustotu. Ak však existuje kontrolná matica s nízkou hustotou, tak kód nazývame *LDPC kód* (z anglického Low Density Parity-check Code). Ak existuje kontrolná matica H s miernou hustotou, kód nazývame *MDPC kódom* (z anglického Moderate Density Parity-check Code). Kontrolná matica MDPC obsahuje viac jednotiek ako matica pre LDPC kód.

Pod pojmom *cyklická matica* chápeme maticu, ktorá vznikne postupnou rotáciou prvého riadku a *kvázi-cyklická matica* pozostáva z blokov takýchto matíc.

Príklad cyklickej matice: $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$

Príklad kvázi-cyklickej matice: $\begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$

Pri QC (kvázi-cyklických) kódach existuje kontrolná matica H zložená z cyklických matíc. Cyklickým posunom v kódovom slove sa vytvorí taktiež kódové slovo. Ak pre LDPC, resp. MDPC kód existuje takáto matica, tak lineárny kód označujeme *QC-LDPC*, resp. *QC-MDPC*.

Hammingova váha vektora predstavuje počet nenulových prvkov vektora. Hammingova vzdialenosť dvoch vektorov predstavuje počet prvkov, ktoré sa odlišujú. Môžeme to chápať aj ako počet nutných zmien na to, aby sa jeden vektor pretransformoval na druhý [12].

2.3 McEliece kryptosystém

McEliece kryptosystém je asymetrický šifrovací algoritmus vyvinutý v roku 1978 profesorom Robertom McElieceom [9]. Bol to prvý kryptosystém svojho druhu, ktorý pri šifrovaní procese využíva náhodnosť. Napriek tomu, že si v kryptografickej komunite nikdy nenašiel príliš veľké prijatie, je jedným z kandidátov na post-quantovú kryptografiu vďaka jeho odolnosti voči útokom, ktoré používajú Shorov algoritmus. Jeho bezpečnosť je založená na probléme dekódovania lineárnych kódov, ktorý je klasifikovaný ako *NP-úplný* problém [3]. Pri jeho konštrukcii je potrebné zvoliť dekódovací algoritmus, ktorý dokáže opraviť t chýb, ktoré sú zanesené do kódu pri šifrovaní. Pôvodný algoritmus využíva tzv. *Goppa kódy* [9] a pozostáva z troch častí: generovanie kľúčov, šifrovanie a deterministický dešifrovací algoritmus. Všetci používatelia takéhoto kryptosystému musia zdieľať rovnaké bezpečnostné parametre n, k, t , kde n je dimenzia priestoru, k je dimenzia kódu (podpriestoru) a t je minimálny počet opravitelných chýb.

2.3.1 Generovanie kľúčov

Alica si zvolí generujúcu maticu G s rozmermi $k \times n$ kódu C , pre ktorý existuje efektívny dekódovací algoritmus. Náhodne vygeneruje maticu S s rozmermi $k \times k$, ktorá nie je singulárna. Náhodne vygeneruje permutačnú maticu P s rozmermi $n \times n$. Vypočíta maticu $G' = SG P$ s rozmermi $k \times n$. Potom jej verejný kľúč je (G', t) a súkromný kľúč sú matice (S, G, P) .

2.3.2 Šifrovanie

Bob chce poslať správu m Alici. Vygeneruje zašifrovanú správu $x = mG' + e$, kde e je náhodný chybový vektor s dĺžkou n , ktorého Hammingova váha nie je väčšia ako t . Táto podmienka zabezpečuje, že kód bude možné pri dešifrovaní dekódovať.

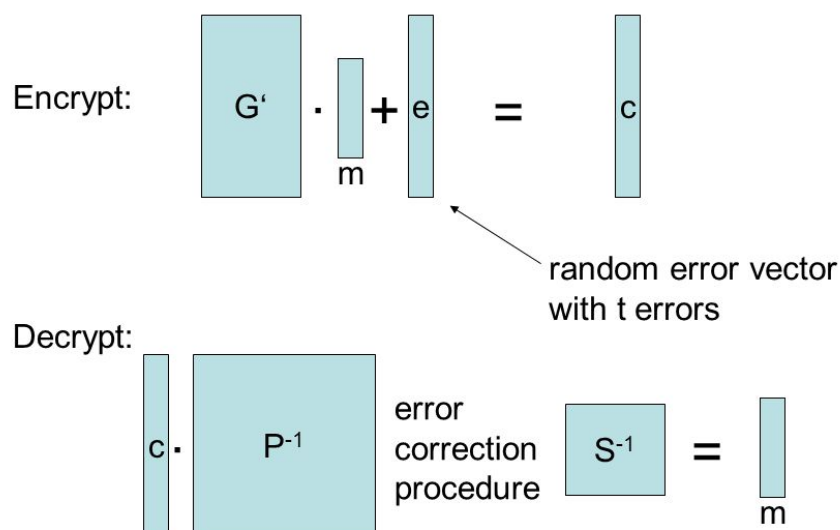
2.3.3 Dešifrovanie

Alica vypočíta inverznú maticu P^{-1} . Ďalej vypočíta xP^{-1} , čo môžeme rozpísať ako:

$$xP^{-1} = (mG' + e)P^{-1} = mSGPP^{-1} + eP^{-1} = mSG + eP^{-1} \quad (3)$$

Na tento výsledok vieme použiť dekódovací algoritmus pre kód generovaný maticou G a teda vieme dostať m' . Posledný krok, ktorý musí urobiť je vypočítať $m = m'S^{-1}$, aby sa dostala k správe, ktorú jej Bob poslal.

The McEliece Cryptosystem



Obrázok 1: Schéma McEliece kryptosystému [5]

McElieceov kryptosystém má viacero výhod. Jednou z nich je napríklad, že šifrovanie a dešifrovanie je rýchlejšie ako pri RSA [4]. Je to hlavne preto, že McEliece vyžaduje iba násobenie vektora a matice, zatiaľ čo RSA umocňovanie čísel. Hlavnou nevýhodou tohto kryptosystému je veľkosť kľúčov. To je dôvod prečo sa tento kryptosystém veľmi neuchytil v praxi. Jednou z mála aplikácií, ktorá používala toto šifrovanie bola aplikácia *Entropy*, čo je decentralizovaná *peer-to-peer* komunikačná sieť, ktorá bola navrhnutá, aby bola odolná proti akejkoľvek cenzúre [10].

2.4 QC-LDPC McEliece

Napriek tomu, že nie je známy praktický útok na McElieceov kryptosystém s Goppa kódmi, mal pôvodný návrh jednu nevýhodu - veľkosť kľúčov. Existuje však spôsob, ako túto veľkosť znížiť a to výber kódov z veľkej automorfnej grupy, akými sú kvázi-cyklické kódy [1]. Avšak mnohé z nich sú napadnuteľné algebraickým útokom, pri ktorom môže útočník vytvoriť systém rovníc a tento systém riešiť pomocou Groebnerových báz. Úspešnosť takéhoto útoku vychádza z vlastností algebraickej štruktúry použitej rodiny kódov. Jedným z dôvodov prečo je takýto útok realizovateľný je, že kvázi-cyklická štruktúra umožňuje drastické zníženie počtu neznámych v systéme [8]. Hrozbe takéhoto útoku môžeme predísť použitím kódov, ktoré nemajú algebraickú štruktúru.

Vhodnými kandidátmi pre tento účel sú LDPC kódy, ktoré nemajú algebraickú štruktúru. Rozdiel voči QC-MDPC kódom spočíva v *redšej* kontrolnej matici H , ktorá umožňuje efektívnu korekciu chýb pri dekódovaní [1].

Ak by chcela Alica dostávať šifrované správy od Boba, potrebuje si vytvoriť verejný kľúč K_V , ktorým Bob bude šifrovať správy adresované Alici a súkromný kľúč K_S , ktorým bude Alica správy dešifrovať. Pre vytvorenie súkromného kľúča si Alica potrebuje zvoliť náhodnú riedku kontrolnú kvázi-cyklickú maticu H a maskovacie matice S a Q . Matice S a Q sú kvázi-cyklické štvorcové invertovateľné matice, pričom S je hustá matica a Q riedka. Tieto tri matice tvoria Alicin súkromný kľúč K_S . Pre vytvorenie verejného kľúča K_V potrebuje Alica z matice H odvodiť generujúcu maticu G . Matice H_1, \dots, H_{n_0} sú riedke cyklické štvorcové matice.

$$H = \begin{bmatrix} H_1 & H_2 & \dots & H_{n_0} \end{bmatrix}$$

Nech matica I je jednotková štvorcová matica, potom generujúcu maticu G vytvorí Alica nasledovne:

$$G = \left[\begin{array}{c|c} I & \begin{matrix} (H_n^{-1}H_1)^T \\ \vdots \\ (H_n^{-1}H_{n-1})^T \end{matrix} \end{array} \right]$$

Po dostaní generujúcej matice môže Alica vypočítať svoj verejný kľúč K_V ako $S^{-1}GQ^{-1}$. Následne môže Bob zašifrovať správu m jej vynásobením Aliciným verejným kľúčom K_V a pridaním chyby e . Pričom Hammingova váha chybového vektoru nie je väčšia ako t , čo je počet chýb, ktorý dokáže kód generovaný maticou G opraviť.

$$x = mK_V + e \quad (4)$$

Alica prijatú správu x dešifruje nasledovne: šifrovanú správu x vynásobí maticou Q , čím dostane $x' = mS^{-1}G + eQ$ (Q je riedka a teda Hammingova váha eQ je malá). x' je kódové slovo QC-LDPC kódu + chyba. Následne použitím vhodného dekódovacieho algoritmu odstráni chyby. Tak dostane správu m' rovnú mS^{-1} , ktorú stačí vynásobiť maticou S a získa pôvodnú správu m .

Správu šifrovanú pomocou QC-LDPC McElieceovho kryptosystému môže Alica dešifrovať podobne ako pri použití QC-MDPC kódov. Alicin verejný kľúč $K_V = S^{-1}GQ^{-1}$ generuje práve QC-MDPC kód. Nech matica $H' = HQ^T$. Keďže H aj Q sú riedke kvázi-cyklické matice, H' je tiež riedka kvázi-cyklická matica, ktorá je kontrolnou maticou pre tento kód, a teda platí $K_V(H')^T = 0$, čo môžeme dokázať nasledovne:

$$\begin{aligned} K_V(H')^T &= K_V(H')^T = \\ &= K_V(HQ^T)^T = \\ &= K_VQH^T = \\ &= S^{-1}GQ^{-1}QH^T = \\ &= S^{-1}GH^T = \\ &= S^{-1}0 = \\ &= 0 \end{aligned} \quad (5)$$

2.5 Bit-flippingové algoritmy

Gallagerov bit-flippingový algoritmus je iteratívny dekódovací algoritmus pre LDPC kódy. V prvom kroku sa vypočíta syndróm správy vynásobením kontrolnej matice H a transponovaného vektoru x^T . Tento výsledok môžeme chápať ako sústavu r rovníc alebo PC rovníc (Parity-check). Ak má rovnica na pravej strane 1, môžeme ju označiť ako UPC (Unsatisfied Parity-check Equation). Následne pre každý bit zo správy x zrátame počet UPC rovníc, do ktorých prispieva. Pre tento výpočet môžeme použiť *Tannerov* graf. Bit, ktorý prispieva do najviac UPC rovníc preklopíme, opäť vypočítame syndróm správy, ale teraz už s aktualizovaným vektorom x' . Ak je výsledok nulový vektor, tak sme správne dekodovali. Ak nie, pokračujeme rovnako v ďalšej iterácii pokým syndróm nie je nulový vektor, alebo sa nedosiahne maximálny počet iterácií. Pre ilustráciu si predvedieme jednoduchý príklad s jednou iteráciou.

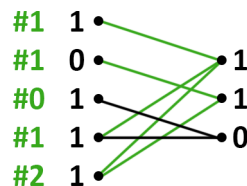
Majme kontrolnú maticu H a vektor x :

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}, x = (1 \ 0 \ 1 \ 1 \ 1)$$

Ďalej si vypočítame syndróm správy s :

$$s = Hx^T = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Pomocou Tannerovho grafu vieme určiť množstvo UPC rovníc, do ktorých jednotlivé bity prispievajú. Na ľavo je transponovaný vektor x a na pravej strane je syndróm správy. Konkrétny bit prispieva do UPC rovnice, ak existuje spojenie v grafe z daného bitu do ktoréhokoľvek bitu syndrómu s hodnotou 1.



Obrázok 2: Tannerov graf na zistenie najpodozrivejšieho bitu

Z Tannerovho grafu vieme určiť najpodozrivejší bit, v našom prípade piaty, nakoľko zasahuje až do dvoch UPC rovníc. Preto ho preklopíme. Dostávame nové x' a opäť môžeme vypočítať syndróm s' :

$$x' = (1 \ 0 \ 1 \ 1 \ 0)$$

$$s' = H(x')^T = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

V tomto prípade je výsledok nulový vektor, čo je ukazovateľom toho, že sme kódové slovo úspešne dekodovali.

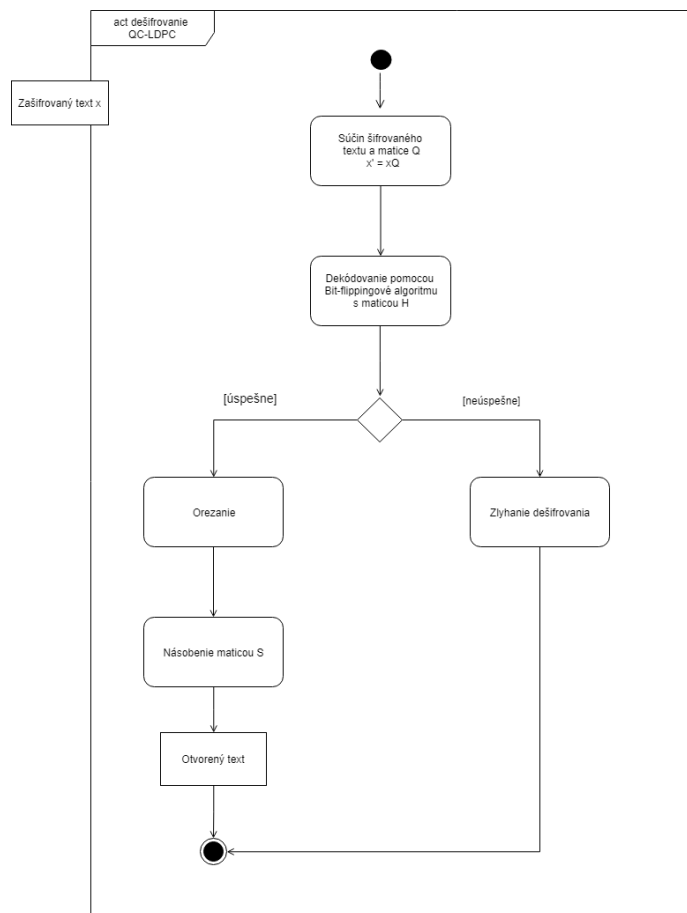
Ako bolo spomenuté vyššie, existujú mnohé modifikácie bit-flippingových algoritmov. Najväčší rozdiel medzi jednotlivými modifikáciami možno pozorovať pri nastavovaní *prahovej hodnoty*, podľa ktorej sa určuje či sa má daný bit preklopiť alebo nie. Vo všeobecnosti, ak je táto hodnota nízka, algoritmus je rýchlejší ale kódové slovo nemusí byť správne dekodované. Naopak ak je táto hodnota vysoká, algoritmus je pomalší, no zväčšuje sa pravdepodobnosť, že povedie k správne výsledku. Preto by sa dalo považovať správne nastavenie tejto hodnoty za kritické.

Modifikáciou bit-flippingového algoritmu môžu byť napríklad, že sa prahová hodnota predpočítava podľa návrhu R. G. Gallagera z jeho práce *Low Density Parity-check Codes* [6]. Ďalší prístup volí túto prahovú hodnotu ako maximálny počet UPC v každej iterácii. Treťou možnosťou je určenie prahovej hodnoty podľa vzorca, v ktorom sa od maximálneho počtu UPC rovníc odpočíta malá hodnota δ . Takýto krok má za následok preklopenie väčšieho počtu bitov na rozdiel od druhej modifikácie, kde sa preklopí jeden alebo len malý počet bitov [7], [8].

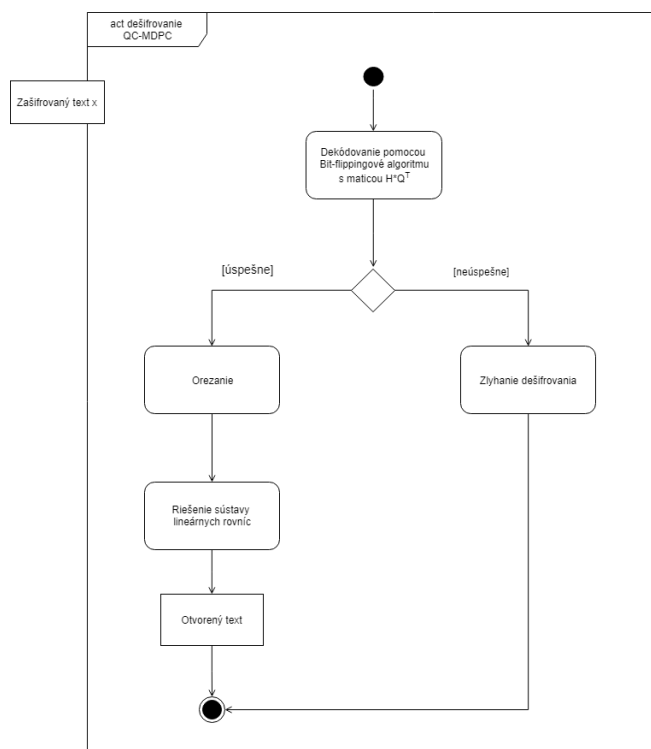
3 Praktická časť

Po naštudovaní problematiky tímového projektu a dostatočnom pochopení teoretickej stránky problému sme začali s praktickou realizáciou práce, na ktorú sme využili knižnicu *BitPunch*. Knižnica je výsledkom niekoľkých diplomových a tímových projektov na Ústave informatiky a matematiky FEI STU. Obsahuje implementáciu McElieceovho kryptosystému s využitím Goppa kódov, ale aj QC-LDPC, resp. QC-MDPC kódov.

Ako sme už v anotácii tímového projektu spomenuli, našou úlohou bolo porovnať efektívnosť dvoch spôsobov dekódovania v McElieceovom QC-LDPC kryptosystéme. Na toto porovnanie sme boli nútení rozšíriť knižnicu BitPunch o ďalšiu funkcionality, konkrétne vynásobenie matice H s maticou Q^T . Následne sme zostrojili implementácie kryptosystémov, pomocou ktorých vykonávame experimenty. UML diagramy implementácií zobrazujeme na obrázkoch nižšie.



Obrázok 3: UML diagram - dešifrovanie v QC-LDPC



Obrázok 4: UML diagram - dešifrovanie v QC-MDPC

Matica H aj matica Q má *kvázi-cyklickú riedku* formu, ktorá je v knižnici BitPunch reprezentovaná pomocou štruktúry s názvom *BPU_T_GF2_Sparse_Qc_Matrix*. Ide o dátový typ, ktorý obsahuje všetky potrebné informácie o matici, ako napr. počet a veľkosť blokov, orientáciu uloženia, ale aj počet riadkov a stĺpcov matice. Samotný riadok matice je uložený ako pole indexov jednotiek. Pri násobení matíc H a Q sme sa čo najviac snažili využiť už implementovanú funkcionalitu knižnice BitPunch. Preto sme najprv konvertovali obe matice na tvar *BPU_T_GF2_Matrix*, vďaka ktorému sme matice jednoduchšie vynásobili. Výsledný produkt násobenia bol následne prevedený späť na pôvodný tvar.

Ako si môžeme všimnúť na UML diagramoch, pri dekódovaní sú použité bit-flippingové algoritmy. Najprv dekódujeme pomocou rýchlejšieho algoritmu s označením *B2*, ktorý má však väčšiu chybovosť. Popis tohto algoritmu môžeme vidieť na obrázku 6. Ak tento algoritmus zlyhá, pokúsime sa dekódovať pomocou algoritmu s označením *A3*, ktorý je zobrazený na obrázku 5. Tento algoritmus je oproti algoritmu *B2* pomalší, ale má nižšiu chybovosť. Vieme ho parametrizovať hodnotou *delta*. V implementácii, ktorú používame je počiatočná hodnota tohto parametra nastavená na 5. V prípade neúspešného dekódovania je táto hodnota postupne znižovaná až na hodnotu nula [7].

Algoritmus Bitflipping A3

Vstup: Kontrolná matica H , zašifrovaná správa c

Výstup: Dekódovaná správa *decoded*, alebo DECODING_FAILURE

```
function DECODE_A3( $H, c$ )  
     $syndrom \leftarrow calculate\_syndrom(H, c)$   
    if  $syndrom = 0$  then  
        return  $c$   
    end if  
    for  $iter \leftarrow 0 \rightarrow max\_iter$  do  
        for  $bit \leftarrow 0 \rightarrow length(c)$  do  
             $upc[bit] \leftarrow calculate\_upc(H[bit], syndrom)$   
        end for  
        for  $bit \leftarrow 0 \rightarrow length(c)$  do  
            if  $upc[bit] \geq max(upc) - \delta$  then  
                Preklop príslušný bit v  $c$   
                 $syndrom \leftarrow calculate\_syndrom(H, c)$   
                if  $syndrom = 0$  then  
                    return  $c$   
                end if  
            end if  
        end for  
    end for  
    return DECODING_FAILURE  
end function
```

Obrázok 5: Bit-flippingový algoritmus A3 [7]

Algoritmus Bitflipping B2

Vstup: Kontrolná matica H , zašifrovaná správa c

Výstup: Dekódovaná správa *decoded*, alebo DECODING_FAILURE

```
function DECODE_B2( $H, c$ )  
     $syndrom \leftarrow calculate\_syndrom(H, c)$   
    if  $syndrom = 0$  then  
        return  $c$   
    end if  
    for  $iter \leftarrow 0 \rightarrow max\_iter$  do  
        for  $bit \leftarrow 0 \rightarrow length(c)$  do  
             $upc \leftarrow calculate\_upc(H[bit], syndrom)$   
            if  $upc \geq b[iter]$  then  
                Preklop príslušný bit v  $c$   
                 $syndrom \leftarrow calculate\_syndrom(H, c)$   
                if  $syndrom = 0$  then  
                    return  $c$   
                end if  
            end if  
        end for  
    end for  
    return DECODING_FAILURE  
end function
```

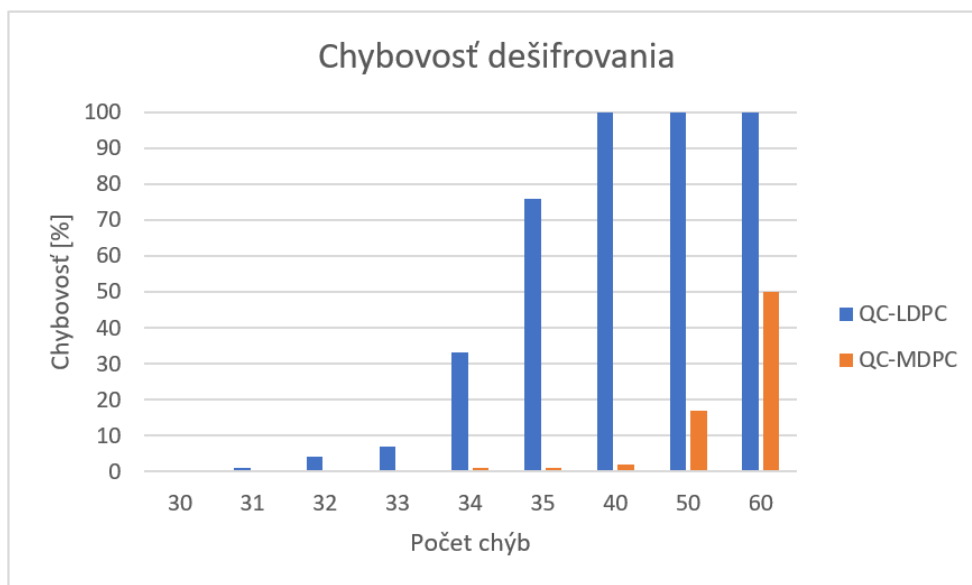
Obrázok 6: Bit-flippingový algoritmus B2 [7]

V nasledovnej časti zobrazujeme výsledky experimentu. Na začiatok sme si vytvorili inštanciu kryptosystému QC-LDPC s nasledovnými parametrami. Počet blokov kvázi-cyklickej kontrolnej matice H bol 3 a jeho riadok mal dĺžku 8192 s hustotou 13. Otvorený text a chybový vektor s vopred určeným počtom chýb sme generovali náhodne, pričom ich dĺžky boli 16 384, resp. 24 576. Proces šifrovania a dešifrovania sme opakovali 100-krát a zaznamenávali sme počet zlyhaní pri danom počte chýb. Počet chýb bol volený v intervale od 30 po 35 s krokom 1 a v intervale 40 až 60 s krokom 10.

Počet chýb	Chybovosť [%]	
	QC-LDPC	QC-MDPC
30	0	0
31	1	0
32	4	0
33	7	0
34	33	1
35	76	1
40	100	2
50	100	17
60	100	50

Obrázok 7: Tabuľka výsledkov experimentu

Z grafu Chybovosti dešifrovania môžeme vidieť, že v prípade použitia QC-LDPC kódu nastáva výrazne vyššia chybovosť ako v prípade QC-MDPC kódu. Ďalej môžeme pozorovať, že od hranice 40 chýb má dešifrovanie pomocou QC-LDPC kódu 100% chybovosť a je teda nepoužiteľný. QC-MDPC má pri tomto počte chýb približne 2% chybovosť a je výrazne lepší. Pri zvyšovaní počtu pridaných chýb môžeme pozorovať nárast chybovosti aj u QC-MDPC.



Obrázok 8: Graf chybovosti dešifrovania

Záver

Cieľom tohto tímového projektu bolo experimentálne porovnať dva spôsoby dešifrovania v QC-LDPC McElieceovom kryptosystéme. Aby sme mohli uskutočniť merania, museli sme najprv doplniť knižnicu BitPunch o novú funkcionálnosť. Zistili sme, že ak pri dekódovaní použijeme maticu, ktorá vznikne súčinom HQ^T , tak dosahuje výrazne menšiu chybovosť ako v prípade použitia matice H . Výsledky zobrazujeme v praktickej časti dokumentácie. Prácou na tomto tímovom projekte sme získali nové vedomosti a praktické skúsenosti z oblasti post-kvantovej kryptografie.

Zoznam použitej literatúry

- [1] BALDI, M., BODRATO, M., AND CHIARALUCE, F. A new analysis of the mceliece cryptosystem based on qc-ldpc codes. In *International Conference on Security and Cryptography for Networks* (2008), Springer, pp. 246–262.
- [2] Berlekamp, Elwyn. *Algebraic coding theory*. World Scientific.
- [3] BERLEKAMP, E., MCELIECE, R., AND VAN TILBORG, H. On the inherent intractability of certain coding problems (corresp.). *IEEE Transactions on Information Theory* 24, 3 (1978), 384–386.
- [4] CANTEAUT, A., AND SENDRIER, N. Cryptanalysis of the original mceliece cryptosystem. In *International Conference on the Theory and Application of Cryptology and Information Security* (1998), Springer, pp. 187–199.
- [5] DOWSLEY, R., VAN DE GRAAF, J., MÜLLER-QUADE, J., AND NASCIMENTO, A. C. Oblivious transfer based on the mceliece assumptions. In *International Conference on Information Theoretic Security* (2008), Springer, pp. 107–117.
- [6] GALLAGER, R. Low-density parity-check codes. *IRE Transactions on information theory* 8, 1 (1962), 21–28.
- [7] Andrej Gulyás. *Implementácia QC-MDPC McEliecovho kryptosystému*.
- [8] GUO, Q., JOHANSSON, T., AND STANKOVSKI, P. A key recovery attack on mdpc with cca security using decoding errors. In *International Conference on the Theory and Application of Cryptology and Information Security* (2016), Springer, pp. 789–815.
- [9] MCELIECE, R. J. A public-key cryptosystem based on algebraic. *Coding Thv 4244* (1978), 114–116.
- [10] MIT Technology Review. *Cryptosystem Resists Quantum Attack*.
- [11] SHOR, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review* 41, 2 (1999), 303–332.
- [12] ZLATOŠ, P. Lineárna algebra a geometria. *Martinus, Bratislava* (2011).

Prílohy

A	Štruktúra elektronického nosiča	II
B	Používateľská príručka	III

A Štruktúra elektronického nosiča

\

\Tímový projekt.pdf

\Zdrojové súbory - *upravená knižnica BitPunch*

B Používateľská príručka

Elektronický nosič obsahuje dokumentáciu k Tímovému projektu a upravenú verziu knižnice BitPunch.

Používateľ musí mať nainštalovaný kompilátor jazyka C. Na kompilovanie knižnice *BitPunch* je možné použiť priložený *makefile*:

- *Kompilovanie knižnice* v príkazovom riadku: v priečinku *../BitPunch/lib* zadať príkaz **make**
- *Spustenie implementácie*: v priečinku *../BitPunch/lib/dist/test* spustiť novovytvorený binárny súbor pomocou príkazu **./BitPunch**