Kód alapú posztkvantum kriptográfia

tudománynépszerűsítő előadás

Nagy Gábor Péter

Szegedi Tudományegyetem
Bolyai Intézet

2022. május 9.

Tagolás

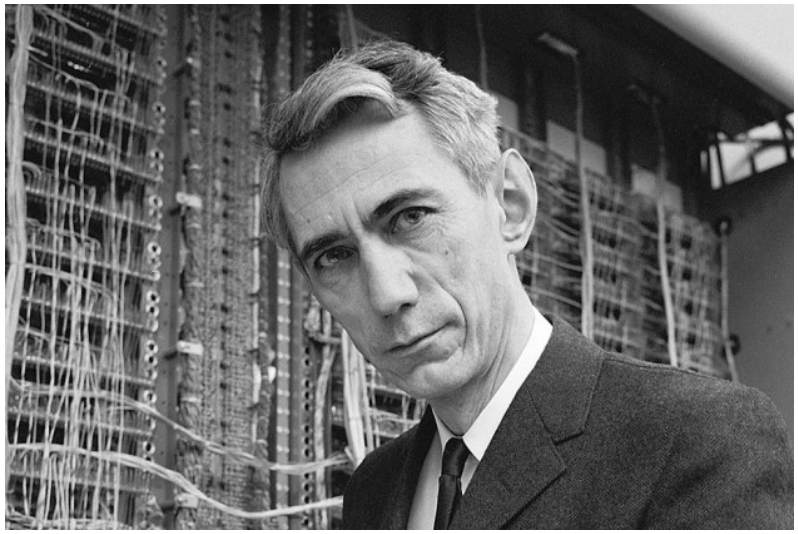
1. Hibajavító kódok
   - Alapfogalmak
   - Dekódolási algoritmusok
   - Résztest részkódok
2. Kód alapú kriptográfia
   - Klasszikus McEliece
   - Kriptoanalízis

Tagolás

1. Hibajavító kódok
   - Alapfogalmak
   - Dekódolási algoritmusok
   - Résztest részkódok
2. Kód alapú kriptográfia
   - Klasszikus McEliece
   - Kriptoanalízis

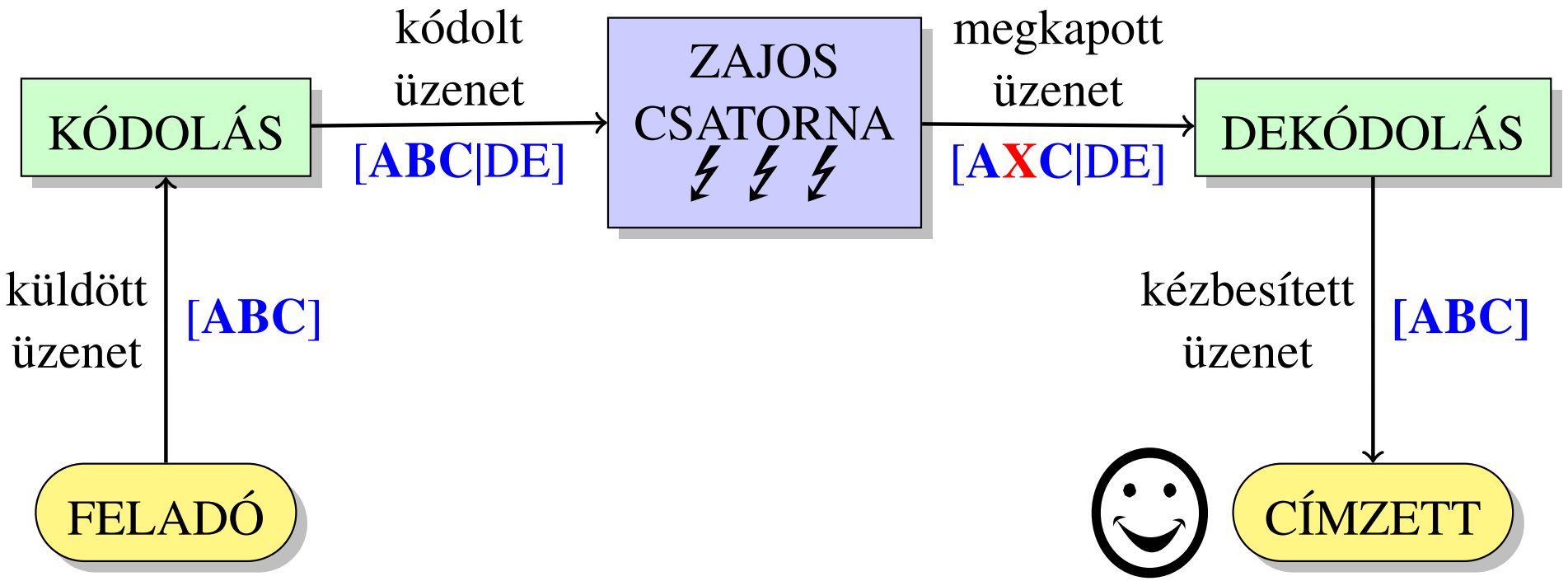
A hibajavító kódok elméletének megalapozása

Richard Hamming
(1915-1998)
amerikai matematikus



Claude Shannon
(1916-2001)
amerikai matematikus

Hibajavítás zajos kommunikációs csatornán



Példa: Hibajavítás a QR-kódban

- A QR-kódokban az $\mathbb{F}_{256}$ véges test felett értelmezett Reed-Solomon kódokat használják.

Lineáris kódok fő paraméterei

A továbbiakban $\mathbb{F}_q$ egy $q = p^m$ rendű véges testet jelöl, ahol p prím.

Definíció: n hosszúságú lineáris kód

- A Q **ábécé** az $\mathbb{F}_q$ véges test.
- $C \leq \mathbb{F}_q^n$ egy lineáris altér.

A főbb paraméterek:

- **Hosszúság** n
- **Dimenzió** k
- **Információs ráta** $R = k/n$
- **Minimum távolság** d

Singleton-korlát

$k + d \leq n + 1$.

A Reed–Solomon-kódok

- Legyen $0 \leq k \leq n \leq q$. Legyenek $\alpha_1, \ldots, \alpha_n$ az $\mathbb{F}_q$ különböző elemei.
- $\mathbf{RS}_k = \{(f(\alpha_1), \ldots, f(\alpha_n) \mid f \in \mathbb{F}_q[X], \deg(f) < k\}$

Tétel

- A Reed–Solomon-kód minimum távolsága $d = n + 1 - k$.
- A **Peterson-algoritmus** ki tud javítani $\lfloor \frac{n-k}{2} \rfloor = \lfloor \frac{d-1}{2} \rfloor$ hibát.

Irving S. Reed (1923-2012)
Gustave Solomon (1930-1996)

W. Wesley Peterson
(1924-2009)

Random kódok

Nincs annál *könnyebb*, mint jó paraméterekkel rendelkező bináris lineáris kódokat készíteni:

Zajos csatornakódolási tétel (Shannon 1948)

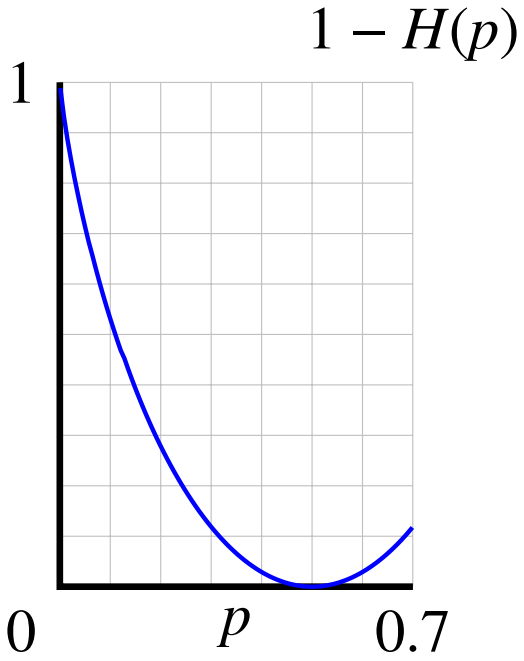
Értelmezzük a bináris entrópiafüggvényt:

$$H(p) = -p \log_2 p - (1-p) \log_2 (1-p).$$

Rögzítsük az $0 < R < 1$ rátát. Ekkor:

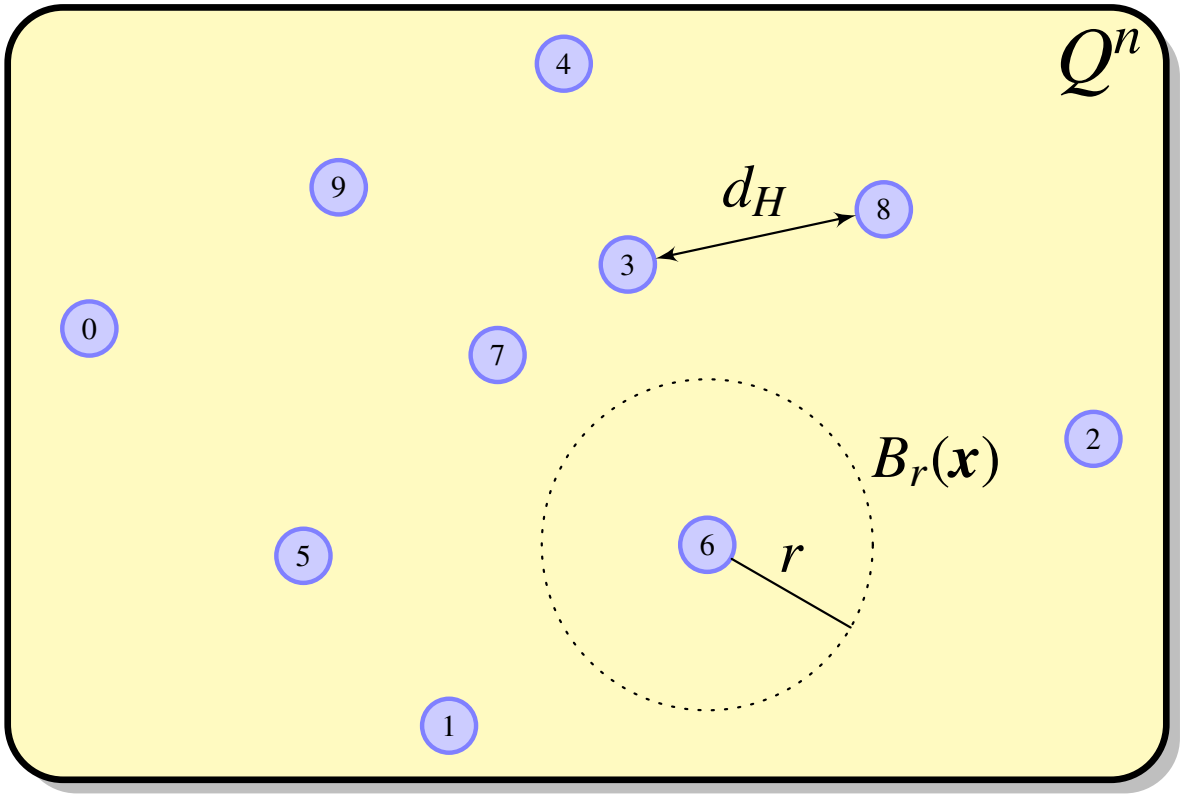
- kellően nagy n esetén
- az n hosszúságú és R rátájú
- **„random”** bináris lineáris kód
- minimum távolsága legalább

$$n \cdot H^{-1}(1-R).$$



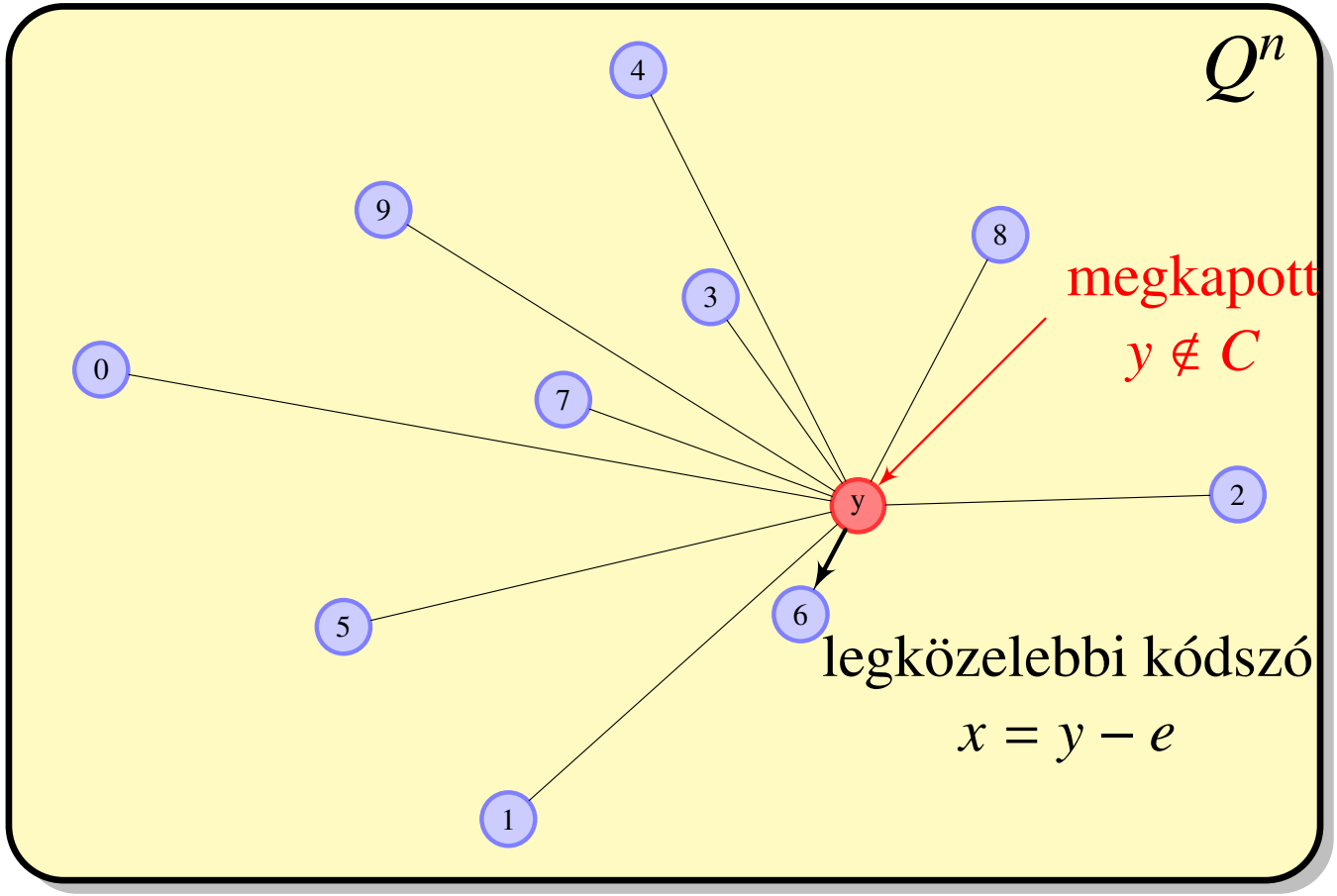
A hibajavító kódolás

- **Véges ábécé** $Q = \{0, 1, \ldots\}$; általában bináris: $Q = \{0, 1\}$
- Egy n **hosszúságú kód** a $C \subseteq Q^n$ halmaz részhalmaza



- Példa: **3-szoros ismétlő kód:** $0 \mapsto 0|00,\ 1 \mapsto 1|11.$
- $Q = \{0, 1\}$, $C = \{000, 111\} \subseteq \{0, 1\}^3$.

Legközelebbi szomszéd *(maximum likelihood)* dekódolás



- $0|00, 1|00, 0|10, 0|01 \mapsto 0|00 \mapsto \mathbf{0}$
  $1|10, 1|01, 0|11, 1|11 \mapsto 1|11 \mapsto \mathbf{1}$.
- Prob(k hiba) > Prob(k-nál több hiba).

A dekódolási feladat

A legközelebbi szomszéd dekódolási feladat

Adott: a $C \subseteq Q^n$ kód és az $\boldsymbol{y} \in Q^n$ vektor.

Keresünk: olyan $\boldsymbol{x} \in C$ kódszót, amire $d_H(\boldsymbol{x}, \boldsymbol{y})$ minimális.

A küszöbértékes dekódolási feladat

Adott: a $C \subseteq Q^n$ kód, az $\boldsymbol{y} \in Q^n$ vektor és a t pozitív egész.

Keresünk: olyan $\boldsymbol{x} \in C$ kódszót, amire

$$d_H(\boldsymbol{x}, \boldsymbol{y}) \leq t.$$

Ha nincs ilyen, akkor az eredmény legyen **„NINCS”.**

- A **küszöbértékes dekódolás** lényegesen könnyebb a legközelebbi szomszéd dekódolásnál, ha $t < d(C)/2$.

A dekódolási feladat lineáris kódokra

Legyen C **lineáris kód** a szokásos paraméterekkel: q, n, k, d.

- C megadható egy $\mathbb{F}_q$ feletti $k \times n$ **generátormátrixszal.**
- A generátormátrix nk elemet tartalmaz, tehát a feladat mérete bitben számolva

$$(nk + n)\log_2(q).$$

- Mivel $k \leq n$, a feladat méretének nagyságrendjét tekinthetjük n^2-nek.

Állítás

- A **kimerítéses keresési eljárás** mindkét dekódolási feladatra exponenciális bonyolultságú megoldási algoritmust ad az n paraméterben.
- Rögzített q mellett a Reed–Solomon-kódok dekódolása *könnyű.*
- A Petersen-algoritmus bonyolultsága n-ben polinomiális.

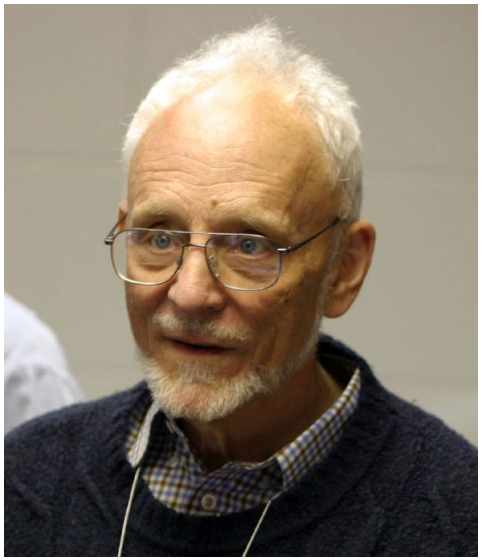
A dekódolás *nehéz* — még a bináris esetben is

Random bináris lineáris kódok *használhatatlanok* a gyakorlatban:

Tétel (Berlekamp, McEliece, van Tilborg, 1978)

Bináris lineáris kódok esetén a küszöbértékes dekódolási feladat ***nagyon nehéz*** (*"NP-teljes"*).

Robert McEliece
(1942-2019)



Elwyn Berlekamp
(1940-2019)

Henk van Tilborg
(1947-)

Az általánosított Reed–Solomon-kód

Definíció: Általánosított Reed–Solomon-kód

Legye q prímhatvány, $0 \leq k \leq n \leq q$. Legyenek $\alpha_1, \ldots, \alpha_n$ az $\mathbb{F}_q$ különböző elemei, $v_, \ldots, v_n$ az $\mathbb{F}_q$ nem nulla elemei.

$$\mathbf{GRS}_k(\boldsymbol{\alpha}, \boldsymbol{v}) = \{(v_1 f(\alpha_1), \ldots, v_n f(\alpha_n)) \mid f \in \mathbb{F}_q[X],\ \deg(f) < k\}.$$

- A *Reed–Solomon* és az *általánosított Reed–Solomon-kódok* paraméterei megegyeznek.
- A Peterson-dekódolás az általános esetben is működik.

Lineáris kódok résztest részkódjai

Definíció: Résztest részkód

Legyen $q = p^m$ prímhatvány és $C \leq \mathbb{F}_q^n$ lineáris kód $\mathbb{F}_q$ felett. Legyen

$$C|_{\mathbb{F}_q} = C \cap \mathbb{F}_p^n.$$

Ekkor $C|_{\mathbb{F}_q} \leq \mathbb{F}_p^n$ *lineáris kód* $\mathbb{F}_p$ felett.

- A résztest részkód minimum távolsága legalább akkora, mint az eredeti kódé.
- Az eredeti kód **dekódolási eljárásai alkalmazhatók** a résztest részkód esetén is
- Küszöbértékes dekódolás esetén a küszöbérték megőrződik.
- A résztest részkód dimenziója

$$\dim(C|_{\mathbb{F}_q}) \geq n - m(n-k).$$

- Nagy m-re pontatlan a becslés, de **nehéz** sokkal jobbat mondani.

Reed–Solomon-alapú bináris kódok

Definíció: Alternáns kód

Az általánosított Reed–Solomon-kódok résztest részkódjait alternáns kódoknak nevezzük.

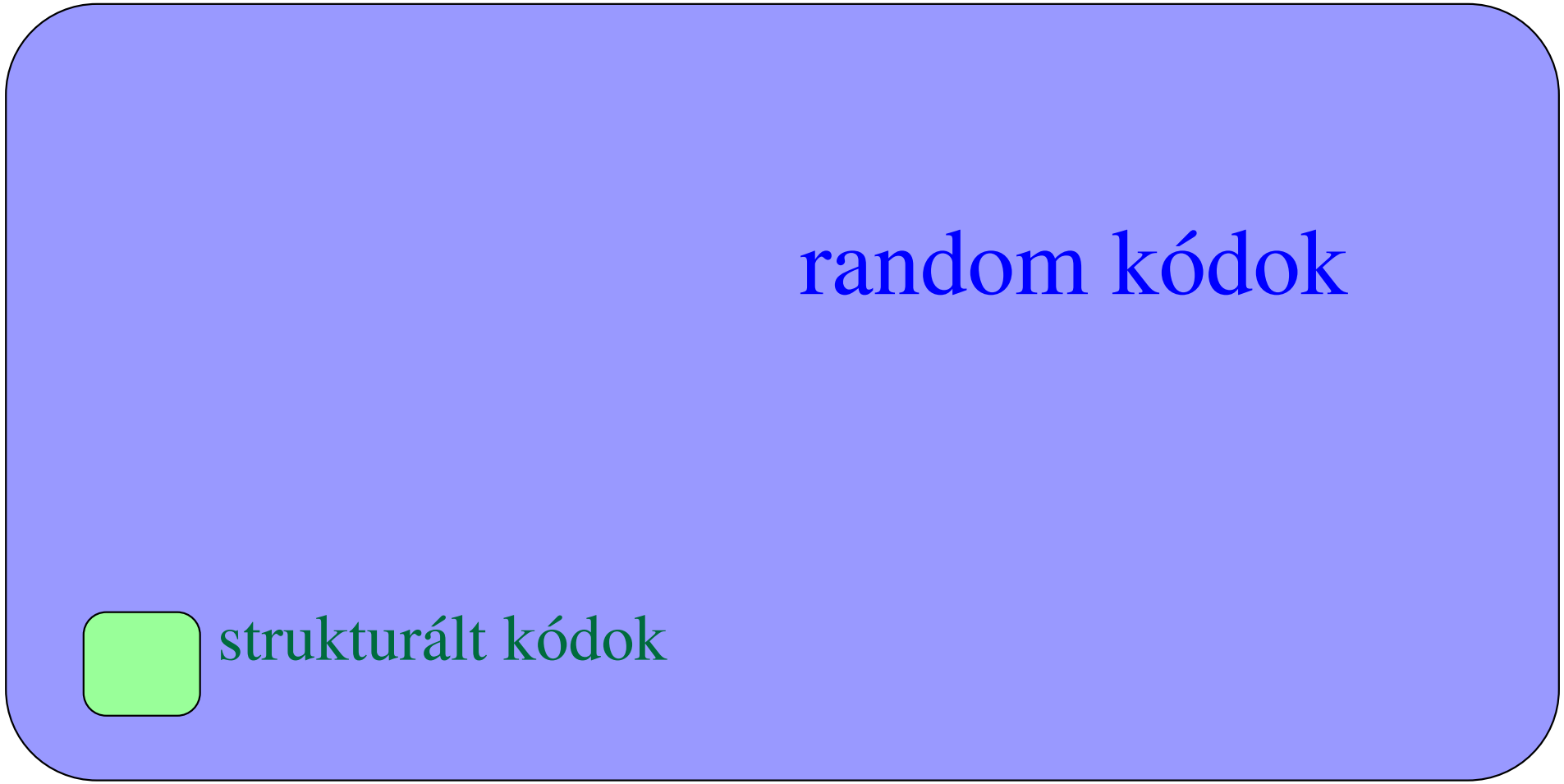
- Az alternáns kód konstrukcióval bináris lineáris kódok egy széles osztályát kapjuk.
- Ezek paraméterei nem különösebben jók.
- Nagy előnyük, hogy ismert hozzájuk polinomiális dekódolási eljárás.
- Tehát megfelelő t küszöbértékkel a dekódolásuk *könnyű*.

Fontos alternáns kód altípusok:

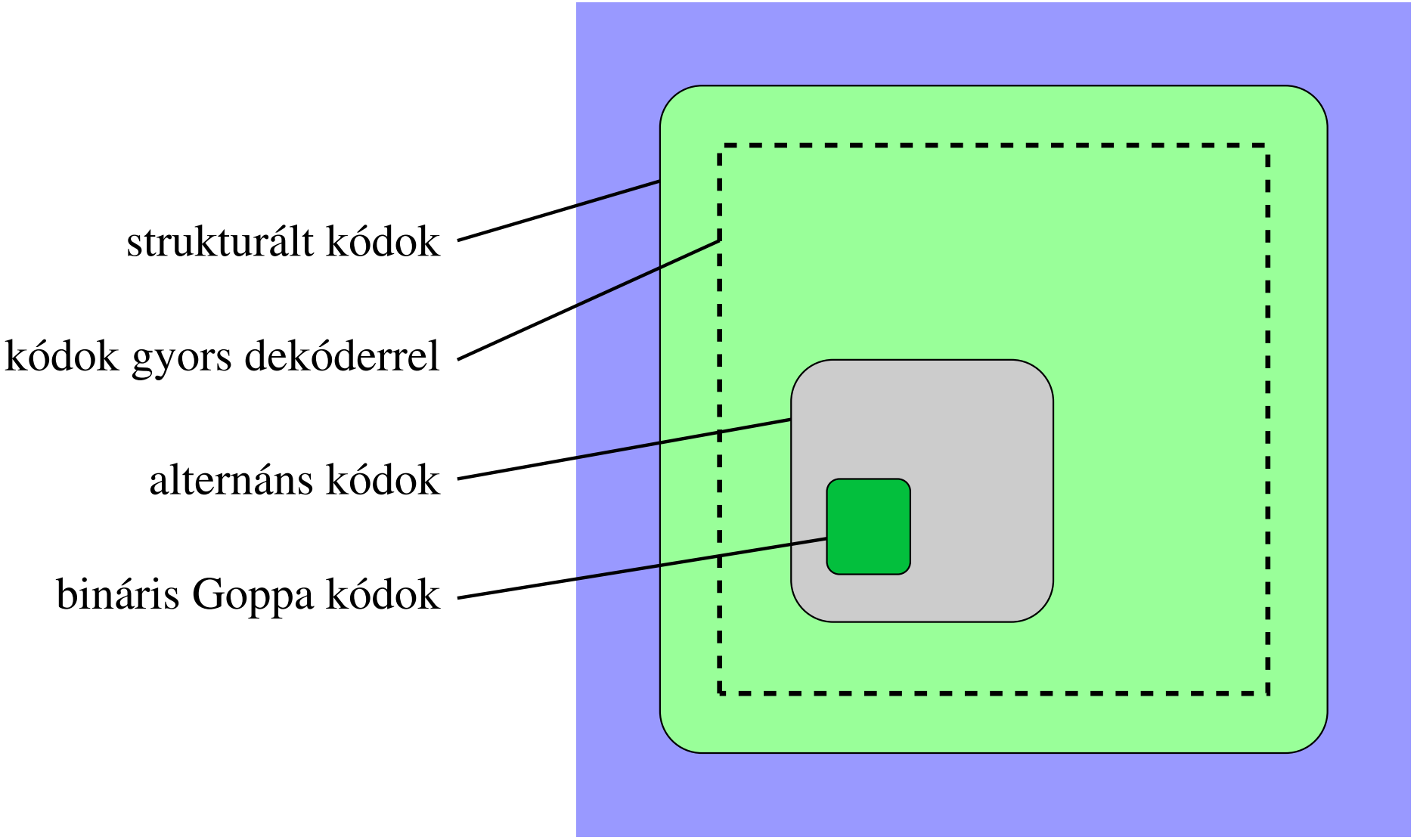
- BCH *(jól skálázható)*
- bináris Goppa *(a várt küszöbérték duplájáig dekódolható)*
- Srivastava

Az n, k, q paraméteres kódok tengere

n hosszúságú, k dimenziós lineáris kódok $\mathbb{F}_q$ felett



A strukturált kódok szigete



Tagolás

1. Hibajavító kódok
   - Alapfogalmak
   - Dekódolási algoritmusok
   - Résztest részkódok

2. Kód alapú kriptográfia
   - Klasszikus McEliece
   - Kriptoanalízis

A McEliece-féle kriptoséma (1978)

Tegyük fel, hogy Alice egy k bites $\boldsymbol{m} \in \mathbb{F}_2^k$ titkos üzenetet akar Bobnak küldeni.

Kulcs generálás

1. Bob választ egy kellően nagy $q = 2^m$ 2-hatványt, $k < n \leq q$ egészt és t küszöbszámot úgy, hogy a megfelelő paraméterekkel rendelkező **alternáns kód** létezik.
2. Bob választ random $\alpha_1, \ldots, \alpha_n$ és $v_1, \ldots, v_n \in \mathbb{F}_q$ elemeket és megkonstruálja a
$$C = \mathbf{GRS}_k(\boldsymbol{\alpha}, \boldsymbol{v}) \cap \mathbb{F}_2^n$$
bináris alternáns kódot.
3. Bob képes t hibát javítani C-ben, mert ismeri az α_i, v_i értékeket. Ezek az értékek alkotják Bob **privát kulcsát.**
4. Legyen G a C egy *generátormátrixa.* Ez Bob **nyilvános kulcsa,** ezt elküldi Alicenak.

Titkosítás és visszafejtés a McEliece-sémában

Titkosítás

1. Alice generál egy **random** t súlyú n hosszú $\boldsymbol{e} \in \mathbb{F}_2^n$ vektort.
2. Alice kiszámítja az
$$\boldsymbol{m}' = \boldsymbol{m}G + \boldsymbol{e}$$
vektort és elküldi Bobnak.

Visszafejtés

1. Bob *(polinomiális) dekódolási algoritmusa* meg tudja határozni azt az **egyetlen** $\boldsymbol{y} \in C$ kódszót, amire
$$d_H(\boldsymbol{m}', \boldsymbol{y}) \leq t.$$
2. Ez pontosan az $\boldsymbol{m}G$ kódszó.
3. Az
$$\boldsymbol{m}G = \boldsymbol{y}$$
lineáris egyenletrendszer megoldásával Bob meghatározza az $\boldsymbol{m}$ üzenetet.

A McEliece-séma biztonsági elemzése *(kriptoanalízis)*

- **Privát kulcs:** $\alpha_1, \ldots, \alpha_n, v_1, \ldots, v_n$. Ezekből konstruáljuk a C alternáns kódot.
- **Nyilvános kulcs:** Az alternáns kód G generátormátrixa.
- A séma biztonságos, ha a nyilvános kulcsból a privát kulcs meghatározása **nehéz feladat.**
- Ennél kicsit többet várunk el: G ne legyen megkülönböztethető egy hasonló méretű **random mátrixtól.**
- Ekkor a biztonságot garantálja, hogy random mátrix esetén a dekódolási feladat **nehéz** (Berlekamp–McEliece–van Tilborg-tétel).
- Ehhez az n, k, t értékeket **kellően nagyra** kell választani.

Klasszikus McEliece modern köntösben

- McEliece eredeti 1978-as javaslata a **bináris Goppa kódokra** épül.
- Ez a $v_1, \ldots, v_n$ *multiplikátoroknak* egy speciális megválasztását jelenti.
- Erre az osztályra a mai napig nem sikerült a privát kulcs visszafejtése.
- Az **eredeti javaslatban** szereplő paraméterek:

$$n = 1014, \quad k = 524, \quad t = 50.$$

- A **2020-as NIST javaslatban** a legmagasabb 256 bites biztonsági szinthez javasolt paraméterek:

$$n = 6686, \quad k = 128, \quad t = 13.$$

- **Túl nagy nyilvános kulcs:** 32, illetve 836 kilobyte.
- *Rengeteg ötlet és javaslat a nyilvános kulcs méretének lényeges csökkentésére – eddig sikertelenül...*