

סיבוכיות

© ארזים

27 ביוני 2017

1 קודים לתיקון שגיאות

תזכורת משיעור שעבר: עבור $\mathbb{F}$ שדה סופי, נאמר שנת קבוצה $C \subseteq \mathbb{F}$ היא קוד. היא קוד לינארי אם C מ"ז מעל $\mathbb{F}$. סימנו $\dim(C) = k$, והגדרנו

$$\text{Rate} = \frac{\log_{|\mathbb{F}|} |C|}{n} = \frac{k}{n}$$

והגדרנו את dist להיות המרחק המינימלי בין שתי מילים בקוד, וכן

$$\text{rel-dist} = \frac{\text{dist}}{n}$$

אמרנו ש- $G \in \mathbb{F}^{n \times k}$ מטריצה יוצרת אם $\text{Im}(G) = C$, $H \in \mathbb{F}^{(n-k) \times n}$ מטריצה בודקת אם $\ker(H) = C$.

דוגמאות

1. קודי ריד-סולומון:

- (א) מרחב ההודעות: מקדמים של פולינום במשתנה אחד ממעלה קטנה מ- k .
- (ב) מילות הקוד: הערכה של הפולינום ב- n נקודות שנקבעו מראש, כלומר עבור v , נשלח $(f_v(\alpha_1), \dots, f_v(\alpha_n))$.
- (ג) מרחק: $n - k + 1$.

2. קודי האדמרד

- (א) מימד: $\log n + 1$.
- (ב) מרחק: $\frac{n}{2}$.

חידה יש n אסירים. לכל אחד יש כובע בצבע לבן/שחור, שנבחר באקראי ע"י הסוהרים. אף אחד לא יודע את צבע הכובע שלו אבל רואה את שאר הכובעים. הסוהר/ת: כולם יוצאו להורג מחר אלא אם כן כולם בו זמנית אומרים מה צבע הכובע שלם וצודקים.

טענה 1.1 אם אחד האסירים/ות למד בקורס זה, וקישר/ה את זה, אז הם יינצלו בהסתברות $1 - \frac{1}{n} \leq$

תרגיל הוכיחו את הטענה. רמז: שאלו את האמינג מה לעשות.

תזכורת למשפט משבוע שעבר: מטריצה $n \times k$ מקרית מעל $\mathbb{F}_2$ נותנת קוד טוב בהסתברות גבוהה.

הוכחנו אותו כך: בחרנו עמודה עמודה את המטריצה היוצרת, וראינו מה ההסתברות שבכל תור אנחנו בוחרים עמודה "טובה" (ששומרת על המרחק המינימלי של הקוד להיות לפחות d). חסמנו את ההסתברות לבחור איפשהו עמודה לא טובה עם חסם האיחוד. קיבלנו שההסתברות לטעות חסומה ע"י

$$\sum_{l=1}^k \frac{2^{l-1}}{2^n} \left(\sum_{i=0}^{d-1} \binom{n}{i} \right)$$

ולכן אם מספר זה קטן מספיק מ-1 נקבל קוד טוב. אם $d < \frac{n}{2}$ אז

$$\sum_{i=0}^d \binom{n}{i} \approx 2^{H(\frac{d}{n})n}$$

כאשר לכל $x \in (0, 1)$ מתקיים

$$H(x) = x \log \frac{1}{x} + (1-x) \log \frac{1}{1-x}$$

לכן אם

$$\frac{2^k \cdot 2^{H(\frac{d}{n})n}}{2^n} < 1$$

אז אנחנו בסדר. זה מתקיים אם $k + H(\frac{d}{n})n < n$. נובע שאם $\frac{k}{n} + H(\frac{d}{n}) < 1$ אז קיימים קודים לינאריים מעל $\mathbb{F}_2$ מממד k ומרחק d . באופן קצת יותר כללי, אם R, δ מקיימים $R + H(\delta) < 1$ אז יש קודים כנ"ל (עם קצב R ומרחק יחסי δ).

1.1 מחיקות בקוד ריד-סולומון

אנחנו עובדים בקוד ריד-סולומון, כלומר $a_0, \dots, a_{k-1}$ מומר ל- $\sum_{i=0}^{k-1} a_i x^i$. $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ קבועים מראש ואנחנו שולחים $(f(\alpha_1), \dots, f(\alpha_n))$. נשים לב שאנחנו צריכים שבשדה יהיו לפחות n איברים.

נניח ששלחנו $\text{Enc}(f)$ והתקבל ווקטור $y \in \mathbb{F}^n$ כך שלכל i ,

$$y_i = \begin{cases} f(\alpha_i) \\ ? \end{cases}$$

הסימן "?" מייצג מחיקה. זהו סוג שגיאה מבחינתנו ונרצה לענות על השאלה הטבעית הבאה:

שאלה עם כמה מחיקות נוכל להתמודד?

טענה 1.2 אם מספר המחיקות קטן או שווה ל- $n - k$ אז אפשר לתקן את ההודעות.

הוכחה: ישנם k ערכים $(f(\alpha_{i_1}), \dots, f(\alpha_{i_k}))$ שקיבלנו ללא מחיקה. ע"י אינטרפולציה ניתן לקבל חזרה את מקדמי f (כי הוא ממעלה $\geq k-1$). דרך אחרת לראות זאת: מטריצת ונרמונדה W היא הפיכה. ניקח את W מסדר k שמורכבת רק מה- k מספרים שכן קיבלנו את הערך של f עבורם. מתקיים

$$W \begin{pmatrix} a_0 \\ \vdots \\ a_{k-1} \end{pmatrix} = \begin{pmatrix} f(\alpha_{i_1}) \\ \vdots \\ f(\alpha_{i_k}) \end{pmatrix}$$

■

ולכן נוכל לחשב את כל ה- a_i .

הערה 1.3 מה שאיפשר את זה זו העובדה שכל k שורות של המטריצה היוצרת של הקוד הן בת"ל.

1.2 טעויות (לא מחיקות)

נניח ששלחנו $\text{Enc}(f)$ והתקבל וקטור $y \in \mathbb{F}^n$ כך ש-

$$\text{dist}(y, \text{Enc}(f)) \leq e < \frac{\text{dist}}{2} = \frac{n-k+1}{2}$$

זו בעיה שונה מבעיית המחיקות. הפעם, אנחנו לא יודעים איזה קואורדינטות שובשו. נשים לב שכאשר אנחנו יודעים איזו קואורדינטות שובשו, אנחנו יכולים להתמודד עם מספר כפול של טעויות.

1.3 אלגוריתם Welsh – Berlekamp

"נדמיין" שהיינו יודעים את מיקום הטעויות: נניח בקואורדינטות $i_1, \dots, i_e$. נגדיר את הפולינום הבא, שנקרא Error Locating Polynomial:

$$E(x) = \prod_{j=1}^e (x - \alpha_{i_j})$$

ונגדיר גם

$$N(x) = E(x) \cdot f(x)$$

אנחנו לא יודעים אותם. מה שאנחנו כן יודעים זה:

$$\deg(f) \leq k-1$$

$$\deg(E) \leq e < \frac{n-k+1}{2}$$

$$\deg(N) = \deg f + \deg E < k-1 + \frac{n-k+1}{2}$$

ונשים לב שלכל $1 \leq i \leq n$ מתקיים:

$$N(\alpha_i) = E(\alpha_i) \cdot y_i$$

כי אם $E(\alpha_i) \neq 0$, אזי $f(\alpha_i) = y_i$. כעת, נחשוב על מקדמי $N(x)$, $E(x)$ כעל נעלמים שרוצים לגלות (כי אם N, E ידועים אזי $f = \frac{N}{E}$). נשים לב שהפולינום מקיימים את המשוואות הלינאריות

$$N(\alpha_i) = E(\alpha_i) y_i$$

לכל $1 \leq i \leq n$. יש לנו n משוואות, וכמות הנעלמים היא

$$\begin{aligned} 1 + \deg E + 1 + \deg N &< 1 + \frac{n-k+1}{2} + 1 + k - 1 + \frac{n-k+1}{2} = \\ &= 1 + k + n - k + 1 = n + 2 \end{aligned}$$

למעשה, זה לא משנה, כי אנחנו יודעים שלמערכת הזו קיים פתרון לא טריוויאלי (N, E) , ולכן נוכל למצוא פתרון לא טריוויאלי כלשהו N', E' .

$$\text{טענה 1.4} \quad \frac{N'}{E'} = \frac{N}{E} = f. \text{ באופן שקול, } N'E = NE'$$

הוכחה: מההגדרה, מתקיים

$$N(\alpha_i) = E(\alpha_i) f(\alpha_i)$$

ממערכת המשוואות אנחנו יודעים שמתקיים

$$\begin{aligned} N \cdot E'(\alpha_i) &= E(\alpha_i) f(\alpha_i) E'(\alpha_i) \\ N'E(\alpha_i) &= E'(\alpha_i) y_i E(\alpha_i) \end{aligned}$$

ואם $E(\alpha_i) = 0$, אזי $y_i \neq f(\alpha_i)$, כלומר שניהם מתאפסים. נחשב דרגות:

$$\begin{aligned} \deg(N'E') &< k - 1 + \frac{n-k+1}{2} + \frac{n-k+1}{2} = n \\ \deg(N'E) &< \dots < n \end{aligned}$$

אבל הפולינומים האלה מסכימים על כל הנקודות α_i , כלומר על n נקודות - לכן הם בהכרח שווים. נחלק באלכסון ונקבל

$$f = \frac{N}{E} = \frac{N'}{E'}$$

■

אם כן יש לנו את האלגוריתם הבא:

אלגוריתם נכתוב מערכת משוואות המקדמים של N, E . נמצא פתרון לא טריוויאלי (N', E') ונחזיר את $f = \frac{N'}{E'}$.

1.4 הרכבה של קודים

נניח שיש לנו קוד מעל שדה גדול $\mathbb{F}_q$, ונניח $q = 2^m$ (אבל זה יעבוד גם מעל כל $q = p^m$, p ראשוני). בהינתן $c_i \in \mathbb{F}_q$, נחשוב על c_i כעל ווקטור באורך m מעל $\mathbb{F}_2$ (למעשה $\mathbb{F}_q$ הוא מרחב ווקטורי מעל $\mathbb{F}_2$ מממד m). אם הקוד המקורי היה מממד k , אורך n , ומרחק d , הקוד החדש (מעל $\mathbb{F}_2$) הוא מאורך $n \cdot m$, מממד $k \cdot m$, אבל המרחק הוא עדיין לכל היותר d . באופן יחסי, R לא השתנה, אבל המרחק היחסי ירד פי m . הודעה בקוד החדש $C_{out} \subseteq \mathbb{F}_q^n \cong \mathbb{F}_2^{mn}$ היא k איברים מעל $\mathbb{F}_q$, או km איברים מעל $\mathbb{F}_2$. נניח בנוסף יש לנו קוד $C_{in} \subseteq \mathbb{F}_2^{n'}$ מממד m .

טענה 1.5 נניח כי C_{out} קוד עם קצב R_{out} ומרחק יחסי δ_{out} , C_{in} קוד עם קצב R_{in} ומרחק יחסי δ_{in} , אזי הקוד $C_{out} \circ C_{in}$ הוא מקצב $R_{out} \cdot R_{in}$, ומרחק יחסי $\delta_{out} \cdot \delta_{in}$. הקידוד עובד כך: בהינתן $k \cdot m$ ביטים, נארוז אותם בתור k איברים של $\mathbb{F}_2^m$. נקודד את האיברים הללו עם C_{out} . כל קוארדינטה של מילת קוד נפתח מחדש כווקטור באורך m מעל $\mathbb{F}_2$. כל קוארדינטה כזו נקודד לפי C_{in} .

הוכחה: ראשית נדבר על הקצב. C_{out} מוציא $\frac{k}{R_{out}}$ קוארדינטות, שהן $\frac{k}{R_{out}}m$ ביטים. כל m מתוכם מקודדים לאורך $\frac{m}{R_{in}}$, ולכן מקבלים בסוף $\frac{mk}{R_{in}R_{out}}$, כשהתחלנו עם mk ביטים. לכן הקצב הוא אכן $R_{out} \cdot R_{in}$.
לגבי המרחק היחסי, אם המרחק האבסולוטי היה d_{out} , d_{in} בהתאמה, אזי מתחילים עם d_{out} קוארדינטות שאינן 0. כל אחת מאלה מקודדת על ידי C_{in} , ובקידוד שלה יש d_{in} קוארדינטות שאינן 0 - בפרט המרחק האבסולוטי הוא לפחות $d_{out} \cdot d_{in}$, ולכן המרחק היחסי הוא לפחות $\delta_{out} \cdot \delta_{in}$. ■

2 חלוקת סוד

נתון "סוד" $\alpha \in \mathbb{F}$. רוצים "לפזר" אותו בין n משתתפים, כך שיתקיים:

- כל k משתתפים יוכלו לשחזר את הסוד.
- אף קבוצה של $d \leq k - 1$ משתתפים לא תדע שום דבר על הסוד.

אלגוריתם נניח שהסוד שייך לשדה בגודל לפחות $n + 1$. נסמנו s_0 . מחלק הסוד עושה את הדבר הבא: מגריל $s_1, \dots, s_{k-1} \in \mathbb{F}$ כלשהו, קובע $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ שאינם אפס (ושונים זה מזה), מגדיר את הפולינום

$$f(x) = \sum_{i=0}^{k-1} s_i x^i$$

השחקן מספר t מקבל את $(f(\alpha_t), \alpha_t)$.