קורס : קודים לתיקון שגיאות ושימושיהם במדעי המחשב
מרצה : ד״ר קלים יפרמנקו
סמסטר : סתיו תשפ״א
תאריך : 22/12/2020

הרצאה 18

List Decoding

מבוא:

בתורת הקודים, קוד לתיקון שגיאות נקרא ניתן לפענוח רשימה (List Decodable) אם בהינתן מילת קוד מורעשת, ניתן לשחזר ממנה רשימה קצרה של מילות קוד אפשריות כך שמובטח שאחת מהן היא מילת הקוד המקורית. זאת בהשוואה למובן הרגיל של פענוח, פענוח יחיד (Unique Decoding), בו משוחזרת מילת קוד אחת שחייבת להיות מילת הקוד המקורית.
פענוח רשימה מאפשר פענוח גם בהינתן מספר גדול יותר של שגיאות מאשר שפענוח יחיד היה מאפשר, שיכול להגיע כמעט עד למרחק הקוד ממש.

הוכחנו שאם ρ הינו רדיוס של list decoding אז אפשר לפענח עם List קבוע אם מתקיים $R \le 1 - h(\rho) - \varepsilon$ ו List יהיה אקספוננציאלי אם $R \ge 1 - h(\rho) + \varepsilon$

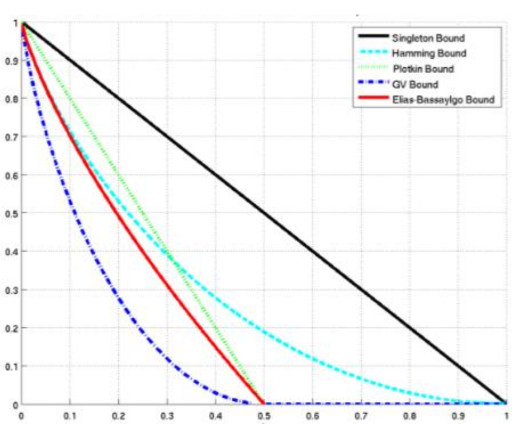
הוכחנו בנוסף שאם מרחק הקוד הינו D אז ניתן לפענח עם List לינארי אם מתקיים ש $\rho = J_2\left(\frac{d}{n}\right)$

Elias Bassalygo bound

משני המשפטים שהוכחנו נסיק שאנחנו יודעים מ J-B שאפשר לפענח עבור $\rho = J_2\left(\frac{d}{n}\right)$, נפעיל על זה את המשפט הראשון ונקבל $R \le 1 - h\left(J_2\left(\frac{d}{2}\right)\right) + \varepsilon$.
כי אחרת אם $R > 1 - h\left(J_2\left(\frac{d}{2}\right)\right) + \varepsilon$ אזי גודלו של List חייב להיות אקספוננציאלי, ובפרט אי אפשר לפענח אם הגודל = 2nd.

קיבלנו משוואה על ה- trade of של הקצב המקסימלי שיכול להיות, והחסם $R \le 1 - h\left(J_2\left(\frac{d}{2}\right)\right) + \varepsilon$ נקרא Elias Bassalygo Bound – חסם על קצב מקסימלי של הקוד עבור מרחק נתון.

נבחין כי זה חסם טוב יותר מ hamming bound ששווה $R \le 1 - h(\frac{d}{2n})$ אך עדיין יש פער בין חסם Elias Bassalygo Bound לחסם הטוב ביותר הידוע – Gilbert Varshamov bound (מלבד חסמים המתקבלים מתוכניות לינאריות)



* אלגוריתם Gilbert Varshamov מגריל מטריצה אקראית על ידי פילוג אחיד ויוצר כך קוד לינארי אקראי.

ובנוסף שיש פער בין חסמים תחתונים ועליונים על הקצב של הקוד עבור מרחק נתון.

*בתמונה המצורפת מתוארים החסמים כתלות המרחק

תכלת - hamming bound

כחול - Gilbert Varshamov bound

אדום - Elias Bassalygo Bound

קורס: קודים לתיקון שגיאות ושימושיהם במדעי המחשב
מרצה: ד"ר קלים יפרמנקו
סמסטר: סתיו תשפ"א
תאריך: 22/12/2020

List – decoding at Reed Solomon

ניזכר בקידוד Reed Solomon ובאלגוריתם unique decoding שראינו בשיעור:

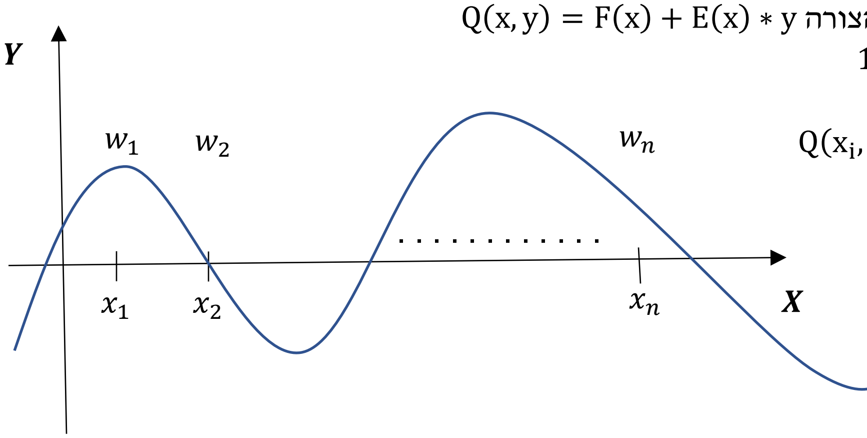
<u>קידוד Reed Solomon:</u>
ההודעה הינה פולינום ומילת הקוד היא הפולינום בנקודות $x_1, .., x_n$

כלומר: $p \longrightarrow p(x_1), p(x_2), \ldots\ , p(x_n)$.

<u>אלגוריתם unique decoding:</u>
<u>קלט:</u> $x_1, \ldots\ , x_n$: n נקודות בהן חישבנו את ערך הפולינום ו- $w_1, w_2, \ldots, w_n$ מילה משובשת.

נמצא פולינום $p(x)$ מדרגה $k-1$ כך שעבור $\left\lceil \frac{n-k}{2} \right\rceil$ נקודות x_i מתקיים $p(x_i) = w_i$. במילים אחרות – פולינום כך שעבור הרבה מהנקודות מתקיים שפולינום בנקודות $x_i = w_i$.

שלב 1 – מחפשים פולינום $Q(x, y)$ מהצורה $Q(x, y) = F(x) + E(x) * y$
כך ש $Q(x_i, w_i) = 0$ לכל $1 \le i \le n$
כאשר $Q(x, y) \neq 0$,
$Q(x_i, w_i) = 0 \ \Rightarrow F(x_i) + w_i E(x_i) \equiv 0$



* אנו יודעים שהפולינום האמיתי עובר דרך הרבה מהנקודות, אנו בונים פולינום בשני משתנים המתאפס בנקודות אלו

לדוגמא:

- אם אין שגיאות, אזי הפולינום $y - p(x)$ מתאפס בנקודות אלה.
  ניתן להבחין זאת על ידי כך שבהצבת w_i נקבל $w_i - p(x_i)$ וכיוון שאין שגיאות $w_i = p(x_i)$ ולכן = 0.

מכפילים בפולינום של השגיאות $E(x)$ ומקבלים לבסוף $Q(x, y) = E(x)(y - p(x))$
נבחין כי אם $E(x_i) = 0$ אז $Q(x_i, y_i) = 0$ גם אם $w_i \neq p(x_i)$

שלב 2 – למצוא את כל הפולינומים מהצורה $y - p(x)$ שמחלקים את $Q(x, y)$

אלגוריתם Sudan

תיאור אלגוריתם

נתון: $w_1, w_2, \dots, w_n \ , x_1, x_2, \dots, x_n$

שלב ראשון של האלגוריתם הוא לבנות פולינום $Q(x,y)$ כך ש: $\deg_x Q = \lceil \sqrt{nk} \rceil$, $\deg_y Q = \left\lceil \sqrt{\frac{n}{k}} \right\rceil + 1$

עבורו מתקיים: $Q(x_i, w_i) = 0$ לכל $1 \le i \le n$

שלב שני הוא לפרק את $Q(x,y)$ לגורמים, ולמצוא בתור רשימה את כל הפולינומים $p(x)$ כך ש $y - p(x)$

הוא גורם של Q.

נבחין כי:

$$y - p_1(x) | \ Q$$

$$y - p_2(x) | \ Q$$

$$\vdots$$

אז:

$$\prod (y - p(x)) | Q$$

ולכן

$$\deg_y \prod_{i=1}^{i=L} (y - p_i(x)) \le \deg_y Q \le \sqrt{\frac{n}{k}}$$

מכאן שגודל הרשימה לכל היותר $\sqrt{\frac{n}{k}}$

הוכחת נכונות:

ראשית נוכיח שקיימים $Q(x,y) \neq 0$.

פולינום כללי מהצורה $\deg_x Q \le l$ ו $\deg_x Q \le r$ נראה באופן הבא:

$$Q(x,y) = \sum_{i=0}^{i=l} \sum_{j=0}^{j=r} a_{ij} x^i y^j$$

a_{ij} הם המשתנים שלנו. נזכור ש $Q(x,y)$ צריך לקיים $Q(x_k, w_k) = 0$, לכן:

$$\sum a_{ij} x_k^i \, w_k^j = 0$$

$x_k^i w_k^j$ ידוע, לכן זו משוואה לינארית על a_{ij}. כיוון ש $k = 1 \dots n$ והתנאי ש $Q(x_k, w_k) = 0$ הוא משוואה לינארית אזי קיימות n משוואות לינאריות.

מספר הנעלמים הוא $n < \sqrt{nk} * \sqrt{\frac{n}{k}} = d_x * d_y$ ולכן בהכרח קיים פתרון השונה מ-0.

ולכן אפשר למצוא מקדמים a_{ij} כך ש $\sum a_{ij} x_k^i w_k^j = 0$ לכל $1 \le k \le n$

קיבלנו כי קיים בהכרח פולינום ולכן השלב הראשון באלגוריתם מתבצע, וכיוון שמציאת איבר בגרעין מטריצה הוא $O(n^3)$ אז זמן הריצה של השלב הראשון באלגוריתם הוא פולינומי.

כעת נוכיח שאם מספר i המקיימים $w_i = p(x)$ הוא לפחות $2\sqrt{nk}$ אז $p(x)$ יהיה ברשימה.
מבניית האלגוריתם אנו יודעים ש $p(x)$ ברשימה אם $y - p(x) | Q(x,y)$

<u>למה:</u> $y - p(x) | Q(x,y) \Leftrightarrow Q(x, p(x)) \equiv 0$
<u>הוכחה:</u> $\Rightarrow$

אם $y - p(x) | Q(x,y)$ אז $Q(x,y) = (y - p(x) * \tilde{Q}(x,y)$
לכן $Q(x, p(x)) = (p(x) - p(x)) * \tilde{Q}(x,y) = 0$

*כיוון שני שקול

<u>טענה:</u> אם מספר i המקיימים $w_i = p(x)$ הוא לפחות $2\sqrt{nk}$, כלומר $2\sqrt{nk}$ מקומות בהם לא היה שיבוש, אז $Q(x, p(x)) \equiv 0$

<u>הוכחה:</u>
ראשית נחשב את דרגת הפולינום $Q(x, p(x))$:

$$Q(x, p(x)) = \sum_{i=1}^{i=d_x} \sum_{j=0}^{j=d_y} a_{ij} x^i p^j(x) \qquad \text{כאשר} \quad d_x = \sqrt{nk},\, d_y = \sqrt{\frac{n}{k}}$$

$$\deg p^j(x) = (k-1) * j$$

$$\deg x^i p^j(x) = i + (k-1) * j < d_x + (k-1) d_y = \sqrt{nk} + k * \sqrt{\frac{n}{k}} = 2\sqrt{nk}$$

ולכן: $\deg (Q(x, p(x)) \le 2\sqrt{nk}$

כעת נוכיח כי $Q(x, p(x))$ הינו פולינום האפס:

מבניית Q מתקיים $Q(x_i, w_i) = 0$. אם $w_i = p(x_i)$ אז $Q(w_i, p(x_i)) = 0$
ולכן לכל x_i עבורו $w_i = p(x_i)$ מתקיים x_i הוא שורש של הפולינום $Q(x, p(x))$ (ניתן להציב ולראות)
ולכן קיבלנו שלפולינום $Q(x, p(x))$ יש $2\sqrt{nk}$ שורשים.
ושמתקיים $\deg(Q(x, p(x)) < 2\sqrt{nk}$
לכן מספר השורשים גדול מדרגת הפולינום ומכאן ש: $Q(x, p(x)) \equiv 0$

לפי הלמה קיבלנו $Q(x,p(x)) \equiv 0$ ולכן $y-p(x)|Q(x,y)$ ולכן $p(x)$ יהיה ברשימה.
בשביל להשתמש בלמה צריך שמספר ה- w_i כך ש $w_i = p(x_i)$ הוא לפחות $2\sqrt{nk}$
ולכן אם היו לכל היותר $n-2\sqrt{nk}$ שגיאות אז המילה הנכונה תהיה ברשימה

ומכאן קיבלנו ש Reed Solomon הוא list decodable עבור $n-2\sqrt{nk}$ שגיאות עם רשימה בגודל $\sqrt{\frac{n}{k}}$

- הרעיון של List Decoding שהצגנו, גרם לפריצת דרך מאוד משמעותית במתמטיקה והשתמשו ברעיון זה בהרבה מקומות: בגאומטריה, קומבינטוריקה ועוד.
  והשימושים הם הרבה מעבר לתורת הקודים.