

סיבוכיות

© ארזים

20 ביוני 2017

היינו באמצע הוכחת המשפט הבא:

משפט 0.1

$$\text{Gap}_v - \text{kCSG}[\delta, 1] \leq_L \text{Gap}_v - \text{k}^l\text{CSG}[\delta^l, 1]$$

הוכחה: הראינו את הבנייה הבאה: נבנה $G' = (V', E')$, כאשר $V' = V^l$ (בלי הגבלת הכלליות, G, G' הם גרפים מלאים). חשבנו על $\Sigma^l = \Sigma'$ - כל צבע הוא וקטור של צבעים, וכל קודקוד הוא וקטור של קודקודים. כאשר צובעים ווקטור עם ווקטור צבעים, בעצם צובעים כל קואורדינטה בצבע המתאים לה. כלומר, $\bar{c}(\bar{v}) = (c_1(v_1), \dots, c_l(v_l))$. הגדרנו את ϕ' באופן שיוודא עקביות של הצביעה - למשל, אם $v \in \bar{u}, \bar{w}$ אז שבשניהם הוא יצבע באותו צבע, וגם שאם $u \in \bar{u}, v \in \bar{v}$ אז נדרוש מהצביעה של $\bar{u}, \bar{v}$ שתקיים את האילוץ של ϕ על (u, v) . ראינו את השלמות בשיעור שעברת וכעת נראה נאותות.

נאותות נראה שאם ניתן לצבוע באופן טוב $\delta^l < \alpha$ מהקודקודים של V' , אזי אפשר לצבוע יותר מאשר δ מהקודקודים של V באופן טוב. נניח, אם כן, שיש צביעה c' הצובעת באופן עקבי יותר מאשר δ^l מקודקודי V' . נגדיר צביעה c באופן הבא: לכל קודקוד $v \in V$ כך שיש $\bar{v} \in V'$ עם $v \in \bar{v}$ שאותו c' צבעה, אז נגדיר את הצבע של v בתור הצבע המושרה מהצביעה c' (הצבע בקואורדינטה המתאימה). בזכות האילוץ הראשון של ϕ' , שדורש קונסיסטנטיות, הצביעה הזו מוגדרת היטב. ברור גם שהצביעה הזו מקיימת את אילוץ ϕ שבהם היא נתקלת, גם בגלל הגדרת ϕ' . נניח שהצלחנו לצבוע כך ε מקודקודי V , כלומר $\varepsilon|V|$ קודקודים. יש $\varepsilon^l|V|^l$ קודקודים של V' שמכילים אך ורק את הקודקודים הללו. נטען כעת שאלה מכילים את $\alpha|V'|$ הקודקודים שאותם c' צבעה. אחרת, היה קודקוד של V' שצבוע ומכיל קודקוד שאינו מתוך $\varepsilon|V|$ הקודקודים שלנו בתוך V , וזה כמובן לא יכול להיות, בגלל איך שהגדרנו את c . מכאן נקבל

$$\delta^l |V|^l < \alpha |V|^l = \alpha |V'| \leq \varepsilon^l |V|^l \\ \delta < \varepsilon$$

כמו שרצינו. לכן סיימנו. ■

מסקנה 0.2 IS קשה לקירוב לכל פקטור קבוע.

הוכחה:

$$\underbrace{\text{Gap}_v - 3\text{CSG}\left(\frac{9}{10}, 1\right)}_{\text{NP-Hard}} \leq_L \text{Gap}_v - 3^l \text{CSG}\left(\left(\frac{9}{10}\right)^l, 1\right) \leq_L \\ \leq_L \text{Gap} - \text{IS}\left[\frac{\left(\frac{9}{10}\right)^l}{3^l}, \frac{1}{3^l}\right]$$

הוכחה: ראינו שזה אומר שאין אלגוריתם קירוב בפקטור $\left(\frac{10}{9}\right)^l$ - אנחנו יכולים לבחור את l כרצוננו, ולכן אין פקטור קירוב קבוע עבור IS. ■

נדבר כעת מעט על הוכחת משפט PCP. מסתכלים על נוסחה $c_1 \wedge \dots \wedge c_m$. או שניתן לספק יחד את כולן, או שניתן לספק לכל היותר $\frac{9}{10}m$ מתוכן. נרצה להראות טרנספורמציה של הבעיה. ■

בעיית Cubic – Solvability (Quadratic – Solvability) הקלט שלנו הוא פולינומים בעלי n משתנים, מדרגה 3, מעל $\mathbb{F}_2$ (אבל אפשר לחשוב על כל שדה סופי קבוע אם רוצים). נסמן את הפולינומים $Q_1(\bar{x}), \dots, Q_m(\bar{x})$. הקלט בשפה אם יש הצבה למשתנים $\bar{\alpha}$ כך שמתקיים $Q_i(\bar{\alpha}) = 0$ לכל i . יש רדוקציה פשוטה מהשפה 3SAT - בהינתן $c_1 \wedge \dots \wedge c_m$, נקודד את בתור $c_i = x \vee y \vee z$

$$Q_i(x, y, z) = (1 - x)(1 - y)(1 - z)$$

. אם יש משתנה עם שלילה, ניקח אותו במקום אחד פחות הוא. השמה מאפסת את Q_i אם ורק אם היא מספקת את c_i . נביט כעת בהשמה $\bar{\alpha}$ ועבורה נביט בווקטור הערכים $(Q_i(\bar{\alpha}))_i$. לכל $\bar{\alpha}$ הווקטור הזה הוא $\bar{0}$ או שיש בו איזושהי קואורדינטה שאינה 0. היינו רוצים שאם יש כזו קואורדינטה, שיהיו הרבה. אנחנו מחפשים טרנספורמציה

$$Q_1, \dots, Q_m \rightarrow Q'_1, \dots, Q'_{m'}$$

כך שאם $\{Q_i\}$ ספיקים אז גם $\{Q'_i\}$ ספיקים, ואם $\{Q_i\}$ אינם ספיקים אז לא ניתן לספק יותר מאשר $\frac{9}{10}m'$ מתוך $\{Q'_i\}$. מסתבר שיש טרנספורמציה לינארית

$$A : \mathbb{F}^m \rightarrow \mathbb{F}^{m'}$$

כך שאם $v \in \mathbb{F}^m$, $\bar{0} \neq v$ אזי בווקטור $A(v)$ לפחות $\frac{m'}{10}$ מהקואורדינטות לא מסתפקות אינן 0. נניח כי $A = (a_{i,j})_{i,j}$. נגדיר

$$Q'_i = \sum_{j=1}^m a_{i,j} Q_j$$

תהי $\bar{\alpha}$ השמה כלשהי. אם לכל i , $Q_i(\bar{\alpha}) = 0$, אזי גם לכל i , $Q_i(\bar{\alpha}) = 0$. אם $\bar{Q}(\bar{\alpha}) \neq \bar{0}$, אזי $\bar{Q}'(\bar{\alpha}) = A(\bar{Q}(\bar{\alpha}))$ מכיל לפחות $\frac{m'}{10}$ קואורדינטות שאינן אפס.

מסקנה 0.3 מהדיון האחרון: בהנחה שיש כאלה מטריצות A , שניתנות לחישוב בזיכרון לוגריתמי, אזי $\text{Gap} - \text{CubicSolvability}\left[\frac{9}{10}, 1\right]$ היא NP-קשה.

1 קודים לתיקון שגיאות

יש לנו הודעה m שעוברת במסלול תקשורת (ערוץ) רועש, ויוצאת ממנו עם רעש, m' . אנחנו רוצים דרך שתאפשר לשחזר את m מתוך m' . למשל, אנחנו מעוניינים בפונקציות Enc, Dec שיאפשרו לנו להתמודד עם 10% טעויות, למשל (כלומר, 10% מהקואורדינטות לכל היותר משתנות). מתברר (ונראה) שניתן לקחת את Enc להיות פונקציה לינארית. אז מה Enc טוב צריך לקיים?

$$\text{Enc} : \mathbb{F}^k \rightarrow \mathbb{F}^n, n > k$$

נסמן dist_H את מטריקת Hamming (כמות קואורדינטות שונות). אנחנו נרצה שלכל $v \neq u$ יתקיים

$$\text{dist}_H(\text{Enc}(u), \text{Enc}(v)) > 2 \frac{n}{10}$$

אם Enc העתקה לינארית, אזי

$$\text{Enc}(v) - \text{Enc}(u) = \text{Enc}(u - v)$$

ולכן

$$\begin{aligned} \text{dist}_H(\text{Enc}(u), \text{Enc}(v)) &= \text{dist}_H(\text{Enc}(u) - \text{Enc}(v), 0) = \\ &= \text{dist}_H(\text{Enc}(u - v), 0) \end{aligned}$$

לכן אם Enc הוא העתקה לינארית הקיימת לכל $v \neq 0$:

$$\text{dist}_H(\text{Enc}(v), 0) > \frac{n}{5}$$

אזי ניתן להתמודד עם 10% טעויות. נגדיר כמה פרמטרים חישוביים:

$$1. \text{ קצב: } R = \frac{k}{n}.$$

$$2. \text{ מרחק יחסי:}$$

$$\delta = \frac{1}{n} \min_{v \neq u} \text{dist}_H(\text{Enc}(v), \text{Enc}(u))$$

ובקוד לינארי גם

$$\delta = \frac{1}{n} \min_{v \neq 0} (\text{Enc}(v), 0)$$

היינו רוצים R גדול ככל האפשר וכן δ גדול ככל האפשר. יש מגבלות על יחס בין שני אלה, למשל בתרגיל הבית נראה כי $R + \delta \leq 1$, ומעל $\mathbb{F}_2$, $R + H(\delta) \leq 1$, כאשר

$$H(\delta) = -\delta \log \delta - (1 - \delta) \log (1 - \delta)$$

הגדרה 1.1 $C \subseteq \mathbb{F}^n$ ייקרא קוד בקצב R ומרחק δ אם $|C| = q^{Rn}$, ולכל $u \neq v \in C$ מתקיים $\text{dist}_H(u, v) \geq \delta n$. C ייקרא קוד לינארי אם הוא תת מרחב ווקטורי של $\mathbb{F}^n$.

הגדרה 1.2 משפחה של קודים $C_n \subseteq \mathbb{F}^n$, $n \rightarrow \infty$, תקרא טובה אם $\text{Rate}(C_n) = R_n$ ו- $\text{dist}(C_n) = \delta_n$ מקיימים

$$\liminf_{n \rightarrow \infty} R_n > 0$$

$$\liminf_{n \rightarrow \infty} \delta_n > 0$$

היינו רוצים לבנות משפחות טובות של קודים עם פונקציות קידוד יעילות ופונקציות פענוח יעילות.

דוגמא לקוד לא טוב - קוד חזרות:

$$(x_1, \dots, x_k) \rightarrow (x_1, x_1, x_1, x_2, x_2, x_2, \dots, x_k, x_k, x_k)$$

נראה כעת קוד טוב מעל שדה גדול.

1.1 קודי Reed – Solomon

יהי $\mathbb{F}$ שדה עם n איברים, והיו $\alpha_1, \dots, \alpha_n$ איברי שדה שונים, ויהי $0 < k < n$. נראה איך לקודד הודעות מאורך k לאורך n באופן שנקבל מרחק $\frac{1}{n}(n - k + 1)$. בהינתן הודעה $\bar{a} = (a_0, \dots, a_{k-1})$, נביט בפולינום

$$f_{\bar{a}}(x) = \sum_{i=0}^{k-1} a_i x^i$$

והקידוד שלנו יהיה

$$\text{Enc}(\bar{a}) = (f_{\bar{a}}(\alpha_1), \dots, f_{\bar{a}}(\alpha_n))$$

הקוד הזה מסומן $\text{RS}(n, k)$. זהו קוד לינארי - באמצעות מטריצת ואנדרמונד. אם $\bar{a} \neq \bar{0}$, אזי בווקטור $\text{Enc}(\bar{a})$ יש לכל היותר $k - 1$ כניסות שהן אפסים, כלומר $\text{dist}(\text{Enc}(\bar{a}), \bar{0}) \geq n - (k - 1) = n - k + 1$ (כי פולינום לא יכול להתאפס ביותר מדי מקומות). בפרט, עבור $k = \frac{n}{2}$ נקבל

$$R = \frac{1}{2}$$

$$\delta = \frac{1}{2} + \frac{1}{n}$$

נראה אפילו אלגוריתם לתיקון של פחות מאשר $\frac{\delta}{2}$ טעויות.

הגדרה 1.3 עבור קוד לינארי מממד k , נניח $C \subseteq \mathbb{F}^n$, מטריצה יוצרת עבור C היא מטריצה $G \in M_{n \times k}(\mathbb{F})$ שמקיימת $\text{Im}(G) = C$.

הגדרה 1.4 עבור קוד לינארי ממימד k , נניח $C \subseteq \mathbb{F}_2^n$, מטריצת בדיקת זוגיות עבור C היא מטריצה $H \in M_{(n-k) \times n}$ שמקיימת $\ker(H) = C$.

דוגמא לקוד פשוט מעל $\mathbb{F}_2$:

$$(a_1, \dots, a_{n-1}) \rightarrow (a_1, \dots, a_{n-1}, a_1 \oplus \dots \oplus a_{n-1})$$

כלומר ביט בדיקת זוגיות מתווסף בסוף. זה קוד ממרחק יחסי $\frac{2}{n}$, ומימדו $n-1$.

$$G = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ 1 & \dots & & 1 \end{pmatrix}$$

$$H = (1 \quad \dots \quad 1)$$

דוגמא קוד Hadamard - ניקח $n = 2^k$, ונחשוב על $[n]$ כעל $\{0, 1\}^k$. ניקח $x \in \{0, 1\}^k$, ונעתיק אותו להיות

$$(\langle x, y \rangle)_{y \in \{0, 1\}^k}$$

כלומר, בקואורדינטה $y \in [n]$ נכתוב את $\langle x, y \rangle$, כשנחשוב עליהם כווקטורים מאורך k מעל $\{0, 1\}$. אפשר להרחיב את קוד האדמארד להכיל גם את $\bar{1} - \text{Enc}(\bar{x})$ לכל $x \in \{0, 1\}^k$. זה נותן קוד מאורך $n = 2^k$ ממימד $k+1$ ומרחק מינימלי $\frac{1}{2}$.

משפט 1.5 קוד לינארי מקרי הוא קוד טוב - יהיו d, k שמקיימים

$$2^k |B(d-1, 0)| = 2^k \sum_{i=0}^{d-1} \binom{n}{i} < \frac{1}{10} 2^n$$

אזי אם נבחר באקראי מטריצה מתוך $M_{n \times k}(\mathbb{F}_2)$ (כל כניסה תהיה 0 או 1 בהסתברות חצי) אז בהסתברות $\frac{1}{2}$ נקבל קוד עם מרחק לפחות d (מעל $\mathbb{F}_2$).

הוכחה: נסמן את עמודות המטריצה שלנו. מה ההסתברות שהווקטור v_1 אינו טוב? אנחנו רוצים לא ליפול בתוך $B(d-1, 0)$, ולכן הסיכוי להיכשל הוא לכל היותר

$$\frac{1}{2^n} \sum_{i=0}^{d-1} \binom{n}{i}$$

בשלב l , כבר בחרנו $l-1$ עמודות. הן מגדירות 2^{l-1} מילות קוד, ונרצה להיות רחוקים מכולן. ההסתברות להיכשל בזה היא לכל היותר

$$\frac{2^{l-1}}{2^n} \sum_{i=0}^{d-1} \binom{n}{i}$$

בסך הכל, ההסתברות לכישלון היא לכל היותר

$$\frac{1}{2^n} \sum_{l=0}^k 2^{l-1} \left(\sum_{i=0}^{d-1} \binom{n}{i} \right) < \frac{2^k}{2^n} \sum_{i=0}^{d-1} \binom{n}{i} < \frac{1}{10}$$

■

וסיימנו.