

הנושא: מתמטיקה "אחרת" - שידור ופיענוח צפנים

הוכן ע"י: ליאורה נוטוב, קרית ים.

מתוך: Salwach C.J., Codes that detect and correct errors, *The College Mathematics Journal*, 19(5), November 1998, p. 402-416.

תקציר: במאמר מציעה הכותבת נושא לשיעור של "מתמטיקה אחרת", בשיעור זה עוסקים בתורת הצפינה, בין השאר נידונים הנושאים יעילות צופן ודוגמאות לצפנים. הכותבת ממליצה לשלב כמה שיעורים שונים משיעורי מתמטיקה רגילים, כדוגמת זה, במהלך שנת הלימודים.

מילות מפתח: צפינה, קודים, פיענוח קודים, הוראת מתמטיקה.

החומר פורסם במסגרת: על"ה 12, הדר תשנ"ג, מרץ 1993, 52-55.

החומר מכיל בנוסף לעמוד הפתיחה: 5 עמודים.

מתמטיקה "אחרת" - שידור ופיענוח צפנים

מדי פעם נשאלת השאלה האם מקבל תלמיד בית ספר, בסוף שתיים עשרה שנות לימוד, מושג אמיתי על מהותה של מתמטיקה. האם תוכנית הלימודים מאפשרת לתלמיד לא רק לדעת פרק מצומצם במתמטיקה, אלא גם לקבל מידע רחב על ענפי המתמטיקה השונים? בתוכנית הלימודים הנוכחית, לצערי, התשובה היא שלילית.

אם נבקש מתלמיד להגדיר מהי מתמטיקה, קרוב לוודאי שיגיד משהו בסגנון: "מתמטיקה היא אוסף מתכונים לפתרון בעיות שונות". כך בהחלט זה נראה מתוכנית הלימודים שלו.

כדי לשנות מצב זה יש לשלב במשך שנת הלימודים לפחות שלוש עד ארבע שעות של מתמטיקה "אחרת". ניתן, למשל, לערוך שיעורים בנושאים כמו: פרקטלים, כאוס, תורת הגרפים, תורת המשחקים, תורת הצפינה וכו'. נושאים אלה נוגעים בחיי היום יום הרבה יותר מהנושאים המוצגים בתוכנית הלימודים הפורמלית - ופותחים לתלמיד צוהר למתמטיקה העכשווית.

במשך שנת הלימודים יש שעות "מתות" שהלימוד הפורמלי בהן אינו אפקטיבי. כך למשל, השיעור הראשון לאחר החזרה מהחופש הגדול, או השבוע האחרון בשנת הלימודים. שעות אלה ניסיתי לנצל כדי להציג נושאים חדשים בכיתות שלי. ברמה של 3 יחידות לימוד קשה מאוד להתחיל לדבר על קיום "מתמטיקה אחרת". גם כאשר עשיתי זאת, יש לציין, לא הצלחתי לעניין את הכיתה כולה אלא רק חלק ממנה.

בכיתות של 4 או 5 יחידות יש התעניינות גדולה יותר, במיוחד אם מסבירים לתלמידים כי הנושא מתפתח בימים אלה ממש. חשוב לציין כי ברגע שהתחלתי לדבר על "מתמטיקה אחרת" ראיתי בעיני התלמידים - בכל הכיתות - הרבה פליאה וחוסר אמון.

במאמר זה אני רוצה להציג נושא לשיעור של "מתמטיקה אחרת": **מבוא לתורת הצפינה (קודים)**.

תורת הצפינה

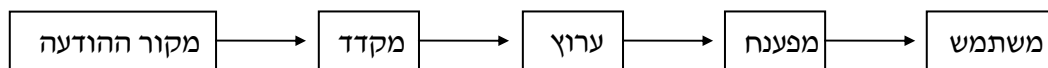
בוא ניזכר במשחק הילדים "טלפון שבור". במשחק זה יושבים המשתתפים זה ליד זה. מטרותו היא להעביר הודעה מהשחקן הראשון לשחקן האחרון. ההודעה מועברת משחקן אחד למשנהו בלחש, כאשר השחקן הראשון משתדל לבחור מילה שקל לשבשה. בסוף המשחק משווים את ההודעה שנשלחה עם זו שהתקבלה.

מתי יצליח השחקן הראשון להטעות את הבאים אחריו? למשל, כאשר יבחר מילה השייכת לקבוצת מילים השונות זו מזו רק באות אחת. לדוגמה: דור. המילים הקרובות יהיו: קור, בור, אור, שור, עור ועוד. אולם אם יבחר השחקן הראשון במילה כמו "שיטפון", קרוב לוודאי כי ההודעה תגיע ללא שיבוש לשחקן האחרון.

גורף נוסף העשוי להביא לשיבוש ההודעה, הוא אורך שרשרת השחקנים. ככל שהשרשרת ארוכה יותר, כך גדל הסיכוי לשיבוש ההודעה. משחקי ילדים זה מדגים את הרעיון של העברת מידע ברשת שידור כלשהי.



אם נגביל את המילים המותרות במשחק לקבוצה מסוימת של מילים, נקבל דוגמה גם לשימוש בצופן ברשת שידור. במקרה זה רשת השידור נראית כך:



זוהי הדמיה מושלמת של רשת שידור אלקטרונית, כאשר הצפנים האלקטרוניים מתבססים על הפרות 0, 1 ומשתמשים בהם על-פי חוקים של אלגברה לינארית מעל שדה בינארי.

הגדרה של צופן בינארי

תהי $F = \{0,1\}$, ותהי V הקבוצה $V = \{x_1, x_2 \dots x_n : x_i \in F, 1 \leq i \leq n\}$

זוהי קבוצת כל ה-n-יות שניתן ליצור מהספרות 0, 1. קבוצה חלקית C של V נקראת **צופן בינארי**, וכל איברי C הם מילות צופן. איברי F הם אותיות האלף בית.

דוגמה: $C = \{0000, 1111\}$

מחיי היום-יום ידוע לנו שלפעמים נופלות טעויות, וההודעה ששידרנו מגיעה למשתמש משובשת; למשל: חייגנו מספר מסוים והגענו למספר אחר. סילוף כזה של ההודעה נובע מרעש הקיים בכל ערוץ תקשורת. כיוון שחשוב לנו שההודעה תגיע בצורה מדויקת, נרצה ליצור צופן שיהיה מסוגל לאתר שגיאה וגם לתקנה במידת האפשר.

דוגמה: בצופן $C = \{0000, 1111\}$ התקבל

א - 0100

ב - 0110

ג - 0000

במקרה א' נפלה כמובן טעות; קרוב לוודאי שההודעה המקורית היתה 0000.

גם במקרה ב' נפלה טעות, אך אין וודאות בקשר להודעה שנשלחה.

במקרה ג' יש סבירות נמוכה ביותר לטעות.

בצופן שבחרנו קל מאוד לאתר טעות ואפילו לתקנה, כיוון שבצופן זה המרחק בין מילות הצופן השונות גדול (בסוף המאמר נסביר ביתר פירוט מהו מרחק בין שתי מילים). הבעיה בצופן הנתון היא ההגבלה על כמות המידע שיש בו. משום כך יש להגדיר דרישות מינימליות מן הצפינה.

הדרישות העיקריות ל תורת הצפינה הן אלה:

1. איתור ותיקון מספר גדול יחסית של טעויות.

2. העברת אינפורמציה רבה ככל האפשר.

3. תהליכים וכללים פשוטים של צפינה ופענוח הודעות.

כיצד משתמשים בצפנים לינאריים לאיתור ותיקון, או לצפינה ופענוח של הודעות מקוריות?

דוגמה טריוויאלית לצופן לינארי היא n-יה של או 1. למשל, 0111001. כל הודעה מוצפנת כזו

מייצגת הודעה מקורית ברורה; אולם במקרה כזה אי אפשר לאתר את הטעות (אם היתה כזו).

אחת השיטות המאפשרות איתור שגיאות ותיקון היא יצירת כפילות תפקידים בהודעה בינארית. נותנים למצפין לשדר כל סיבית (אות בינארית) של ההודעה המקורית מספר מוסכם של פעמים, R . אוסף כל ההודעות המוצפנות האפשריות מסוג זה נקרא **צופן החוזר על עצמו**. ההודעה המתקבלת, העשויה להיות מסולפת, ניתנת לפענוח באמצעות החלפת כל R -יה בסיבית יחידה - זו המופיעה בשכיחות הגדולה ביותר (במקרה של R זוגי, אם R -יה המתקבלת מספר שווה של 0 ו-1, נקבע כי הסיבית תהיה 0).

לדוגמה: בטבלה הבאה $k = 4$ ו- $R = 3$, (k הוא מספר האותיות בהודעה המקורית).

ההודעה המקורית	ההודעה המוצפנת	ההודעה המתקבלת	ההודעה המפוענחת
0 1 1 0	000 111 111 000	000 111 101 000	0 1 1 0
0 1 1 0	000 111 111 000	000 011 100 000	0 1 0 0
0 1 1 0	000 111 111 000	000 011 101 101	0 1 1 0

כפי שרואים בטבלה, אפשר תמיד לאתר עד 2 טעויות בשלשה אך השגיאה ניתנת לתיקון רק אם בשלשה הנתונה שובשה ספרה אחת בלבד. ואומנם, בשורה השניה אנו רואים מקרה שבו מופיעות שתי טעויות באותה שלשה בהודעה המתקבלת. במקרה זה אי אפשר לאתר את הטעות ואף לא לתקנה כי שתי השגיאות מופיעות באותה שלשה. באופן כללי, בצפנים החוזרים על עצמם אפשר תמיד לאתר נוכחות של עד $[R - 1]$ טעויות בכל R -יה. ה- R -יה ניתנת לתיקון אם מופיעות בה עד

¹ $[x]$ מסמן את המספר השלם הקטן או שווה ל- x .

$\frac{R-1}{2}$ טעויות. אם מופיעות יותר מ- $\frac{R}{2}$ טעויות ב- R-יה, יתקבל תמיד ביטוי שגוי בצורת פענוח זר.

מהי יעילות של צופן? כדי לחשב זאת נסמן ב- p את ההסתברות שהסיבית (0 או 1) המשודרת בערוץ תתקבל באופן שגוי. $q = 1 - p$ היא ההסתברות שהסיבית המשודרת והסיבית הנקלטת תהינה שוות.

נחשב את ההסתברות לקבל שדר באורך $k \cdot R$ בלא שגיאות (שוב, k הוא מספר האותיות ו- R מספר החזרות על כל אות).

לדוגמה, $R = 5$ פירושו שכל סיבית משודרת חמש פעמים. לפי האמור לעיל, אפשר לאתר ולתקן עד שתי שגיאות בכל חמישיה. ההסתברות שלא תהיה אף שגיאה בחמישיה אחת היא q^5 ; ההסתברות שתהיה שגיאה אחת היא $5q^4p$; וההסתברות לקבל שתי שגיאות היא $\binom{5}{2}q^3p^2$. לכן, ההסתברות לקבל לכל היותר שתי שגיאות בחמישיות הנקלטות היא $(q^5 + 5q^4p + 10q^3p^2)^k$. בטבלה הבאה מחושבת ההסתברות לקבל שדר נכון כאשר $R = 1, 3, 5$, $p = 0.05$, $k = 4, 6$.

R \ k	1 q^k	3 $(q^3 + 3q^2 \cdot p)^k$	5 $(q^5 + 5q^4p + \binom{5}{2}q^3p^2)^k$
4	0.8145	0.9713	0.9954
6	0.7573	0.9573	0.9931

למרות ההסתברות הגבוהה לקבלת הודעות נכונות על ידי שימוש בצופן זה, הצופן החוזר על עצמו אינו יעיל במיוחד. כל הודעה מוכפלת פי R ועצם שידורה של הודעה ארוכה כזו מגביר בהרבה את ההסתברות להיווצרות רעש בערוץ (הדבר דומה לשרשרת משתתפים ארוכה במשחק "טלפון שבור"). כמו כן, שידור הודעות ארוכות גורם לעומס רב על קו השידור.

צפנים של ביקורות זוגיות (parity checks binary codes)

אנו מחפשים צפנים בינאריים שיאתרו ויתקנו שגיאות, אך יהיו קצרים מן הצפנים החוזרים על עצמם.

א. צופן בדיקות שוויון כולל

$x = x_1x_2 \dots x_{n-1} \in F^{n-1}$; $F = \{0,1\}$ מגדירים את x_n כסכום (מודולו 2) של $n-1$ הסיביות המרכיבות

את x $\left(x_n = \sum_{i=1}^{n-1} x_i\right)$. במקרה זה תמיד יתקיים $\sum_{i=1}^n x_i = 0 \pmod{2}$ קבוצת כל ה-n-יות מהצורה

$$V = \left[x_1x_2 \dots x_{n-1} \mid \sum_{i=1}^{n-1} x_i \mid x_i \in F \ 1 \leq i \leq n \right] \text{ נקראת צופן ביקורת זוגיות כולל בינארי } (n, n-1).$$

נניח שההודעה המקורית מוצפנת כ- $x_1x_2 \dots x_n$ משודרת ונקלטת כ- $\hat{x}_1 \hat{x}_2 \dots \hat{x}_n$. אם $\sum_{i=1}^n \hat{x}_i \neq 0$

פירוש הדבר שיש בוודאות מספר אי זוגי של טעויות. שיטה זו יעילה הרבה יותר **באיתור** שגיאה בודדת מצופן החוזר על עצמו שבו $R = 2$. כלומר, כאשר מכפילים פי שניים את אורך ההודעה על ידי שידור כל סיבית פעמיים.

ב. צופן בדיקות שוויון

נכליל את רעיון בדיקות השוויון וננסה למצוא שיטה שלא רק תאתר שגיאות, אלא אף תתקן. ראינו כי לצופן בינארי (n, k) של C מעל F (כאשר k הוא אורך ההודעה המקורית ו- n - אורך המלה המוצפנת) יש 2^k מלות צופן (או וקטורים). כעת נראה כי לצופן כזה יש k מצבי מידע ו- $n - k$ מצבי בדיקות שוויון. נבחר בסיס ל- C ונמספר את הוקטורים שלו מ- 1 ועד k. המטריצה G, ששורותיה הם וקטורי הבסיס C, נקראת **המטריצה היוצרת** של C משום שהצופן C בנוי מכל הקומבינציות הלינאריות של שורות G. ביצוע פעולות אלמנטריות על השורות או אולי פרמוטציה של העמודות (הנותנת צופן שקול לצופן C) יכול להביא למטריצה יוצרת G שצורתה $G = (I_k | A)$,

כאשר I_k היא מטריצת יחידה מסדר $k \cdot k$ ו- A - מטריצה מסדר $k \cdot (n - k)$ (אם משנים את סדר העמודות כדי לקבל את G אזי יש לשנות באותו אופן את הרכיבים המתאימים של ההודעה המוצפנת).

דוגמה 1: תהי מטריצה יוצרת לצופן בינארי $C(5,3)$

$$G = \left(\begin{array}{ccc|cc} 100 & 11 \\ 010 & 10 \\ 001 & 11 \end{array} \right)$$

$$\left(A = \begin{pmatrix} 11 \\ 10 \\ 11 \end{pmatrix} \quad \text{כאן} \right)$$

כל הודעה $x_1x_2x_3 \in F^3$, כאשר $F = \{0,1\}$ מוצפנת לוקטור 5 - מימדי על ידי הטרינספורמציה

$$(x_1, x_2, x_3, x_4, x_5) = (x_1, x_2, x_3) G =$$

$$(x_1, x_2, x_3, x_1 + x_2 + x_3, x_1 + x_3)$$

$$(x_1, x_2, x_3, x_1 + x_2 + x_3, x_1 + x_3) \in F^5 \quad \text{ברור כי:}$$

$$x_4 = x_1 + x_2 + x_3 \quad \text{לפיכך}$$

$$x_5 = x_1 + x_3$$

שתי משוואות אלה הן $n - k = 2$ בדיקות זוגיות (אחת ל- x_{3+1} והשניה ל- x_{3+2}). במילים אחרות, כדי ש- $x_1x_2x_3x_4x_5$ תהיה הודעה מוצפנת צריכות להתקיים שתי המשוואות

$$(-x_1) + (-x_2) + (-x_3) + x_4 = 0$$

$$(-x_1) + (-x_3) + x_5 = 0$$

ובצורה מטריציאלית

$$H \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -1 & -1 & -1 & 10 \\ -1 & 0 & -1 & 01 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

השוויון הוא מודולו 2 (יש לזכור כי $-1 = 1 \pmod{2}$).

המטריצה H היא מסדר $2 \cdot 5$ וצורתה $H = (-A^t | I_2)$ (א' היא המטריצה המוחלפת של A).

באופן כללי הודעה $x_1x_2x_3 \dots x_k \in F^k$ ניתנת להצפנה כ- n יה של C ב- F^n , באמצעות חישוב G

$(x_1x_2 \dots x_k)$, כאשר G מטריצה יוצרת לצופן $C(n,k)$.

אם לצופן (n, k) יש מטריצה יוצרת מהצורה $G = (I_k | A)$, אזי המטריצה $H = (-A^t | I_{n-k})$ נקראת מטריצת בדיקות השוויון של C . בדוגמה שלפנינו מטריצת ביקורת הזוגיות היא

$$H = \begin{pmatrix} -1 & -1 & -1 & 10 \\ -1 & 0 & -1 & 01 \end{pmatrix}$$

נראה בדוגמה כי צופן בדיקות ביקורת זוגיות, כולל בינארי $(n, n - 1)$, הוא מקרה פרטי של צופן בדיקות זוגיות.

דוגמה 2: נתבונן במקרה $n = 4$ של צופן בדיקות שוויון כולל (4,3). מטריצה יוצרת G לצופן זה תהיה

$$G = \begin{pmatrix} 1001 \\ 0101 \\ 0011 \end{pmatrix}$$

שכן דרשנו ש- $x_n = x_1 + x_2 + \dots + x_{n-1}$

ואכן $(x_1, x_2, x_3) G = (x_1, x_2, x_3, x_1 + x_2, x_3 + x_4)$

מטריצות ביקורת הזוגיות במקרה זה היא $H = (-A^t | I_1) = (-1, -1, -1 | 1) = (1, 1, 1, 1)$ נזכור שוב: $1 \pmod{2} = 1$

כלומר צריך להתקיים השוויון $H(x_1, x_2, x_3, x_4)^t = 0$ שפירושו $x_1 + x_2 + x_3 + x_4 = 0$, וזו התוצאה שקיבלנו ביקורת זוגיות כולל.

כדי לבדוק כמה שגיאות שידור מסוגל צופן זה לאתר ולתקן, נצטרך להגדיר מהו משקלה של מלת צופן ומהו מרחק בין שתי מלות צופן.

משקל של מלת צופן הוא: $wt(x) = |\{i : x_i \neq 0\}|$

כלומר, מספר הרכיבים השונים מאפס.

למשל, משקל המילה $x = 10110$ הוא $wt(10110) = 3$

מרחק בין שתי מילים x ו- y , $x, y \in C$ הוא $D(x, y) = |\{i : x_i \neq y_i\}|$ כלומר, מספר הרכיבים שבהם שונות המילים זו מזו.

קל להראות כי פונקצית המרחק D היא מטריקה של הצופן C , כלומר, עבור $x, y, z \in C$ מתקיים:

- (1) $D(x, x) = 0$
- (2) $D(x, y) = D(y, x)$
- (3) $D(x, z) \leq D(x, y) + D(y, z)$

כמו כן קל לאמת, כי $D(x, y) = wt(x - y)$

לכן מינימום המרחק בין שתי מילים שונות בצופן $d = \min \{D(x, y) \mid x \neq y, x, y \in C\}$ שווה למינימום המשקל השונה מאפס של מילות הצופן.

לדוגמה, ניקח צופן ביקורת זוגיות כולל בינארי $(n, n-1)$, $d = 2$, המתאר שגיאה אחת בלבד. ככל שמינימום המרחק בין המילים בצופן גדול יותר, מלות הצופן כמובן רחוקות יותר זו מזו, ויש להניח שניתן לאתר ולתקן שגיאות שידור רבות יותר. זהו תוכנו של המשפט הבא.

משפט²: יהי d המשקל המינימלי השונה מ-0 של צופן C , אזי:

1. צופן C יכול לאתר מ-1 עד $d-1$ שגיאות שידור;
2. לפי תיאורית הפענוח של סבירות מקסימלית³ אפשר לתקן $t = \lfloor d - \frac{1}{2} \rfloor$ שגיאות בהודעות משודרות.

הכרנו את הרעיון שלפיו בנויים צפנים של בדיקות שוויון. צפנים מסוג זה יעילים ושימושיים. בין כל הצפנים לבדיקות שוויון קיימת קבוצה שימושית ומוכרת ביותר והיא קבוצת הצפנים המושלמים של המינג (Hamming). צפנים אלה הם מסוג (7,4), ועליהם לא נרחיב כאן את הדיבור.

במאמר זה ראינו שתי דוגמאות פשוטות לצפנים. תורת הצפנים התפתחה בסוף שנות ה-40 עם פיתוחה של האלקטרוניקה ומאז התרחבה לכדי מדע לממש תוך כדי שימוש רב ביותר בתחומים שונים של מתמטיקה: חבורות וחוגים, אלגברה לינארית ומודרנית, קומבינטוריקה, תורת המספרים ועוד. מדע זה נמצא עדיין בשלבי פיתוח ויש לשער שהו עדיין רחוק ממיצוי.

² במאמר זה נסתפק בציטוט המשפט ולא נוכיח אותו.

³ פענוח בסבירות מקסימלית (maximum likelihood decoding) משמעו שכאשר מקבלים הודעה $\hat{x} = \hat{x}_1 \hat{x}_2 \dots \hat{x}_n$

שהיא שגויה בוודאות יש לפענח אותה להודעה $z = z_1 z_2 \dots z_k$ באופן z יהיה שונה מ- $\hat{x}$ במינימום של מקומות וכאשר $z \in C$. במאמר זה לא נתאר את שיטת הפענוח ולא נענה על אלת יחידות הפענוח.