סיכומים למבחן בקורס יסודות הקריפטוגרפיה

פרופ' בני שור, סמסטר א' תש"ע (2009-2010)

הגדרות וסימנים, הקדמה:

- פונקציית הצפנה: E.
- פונקציית פענוח: D.
- מפתחות (הצפנה, פענוח): k_1, k_2 (במערכת הצפנה סימטרית $k_1 = k_2$).
- מרחב ההודעות: M.
- דרישת קונסיסטנטיות: לכל $m \in M$ ומפתחות הצפנה/פענוח תואמים k_1, k_2 מתקיים: $D_{k_2}\left(E_{k_1}(m)\right) = m$.
- Plaintext: ההודעה לפני ההצפנה.
- Ciphertext: ההודעה המוצפנת: $E_{k_1}(m)$.

מודל התקשורת:

- אליס ובוס שני צדדים המנהלים תקשורת מוצפנת על קו תקשורת אמין באמצעות אותו אלגוריתם הצפנה.
- אלגוריתמי ההצפנה והפענוח ומרחב ההודעות ידועים לכל: E,D,M.
- חווה (Eve) היא מאזינה לקו התקשורת ויודעת גם את $E_{k_1}(m)$, אינה יודעת את k_1, k_2 ורוצה לגלות את m. היריב מוגבל חישובית (לרוב נניח כי מוגבל פולינומיאלית).
- מטרות ההצפנה: אף יריב לא יוכל לגלות את m או שום אינפורמציה חלקית בעלת משמעות על m.

דוגמאות להצפנות פשוטות:

- Shift cipher: הזזת התווים במרחק קבוע. למשל "בית" יעבור בהזזה ב-1 ל-"גכא". חסרון: מרחב המפתחות קטן מדי.
- Substitution cipher: צופן הצבה, פרמוטציה כלשהי על האותיות. גודל מרחב המפתחות (בשפה האנגלית): $26! \approx 4 \cdot 10^{27}$. מעבר על כל התמורות האפשריות בלתי אפשרי, אך ניתן להשתמש בסטטיסטיקות התפלגות אותיות / זוגות / שלשות בשפה כדי לפענח.

מרחב מפתחות גדול הינו תנאי הכרחי אך לא מספיק בשביל בטיחות ההצפנה.

Perfect cipher:

- יהי מרחב ההודעות $M = \{0,1\}^n$.
- בהינתן ciphertext שנסמנו c, הסיכוי ש-$D_{k_2}(c) = m$ לכל הודעה $m \in M$ שווה לסיכוי ללא ידיעת c ש-m היא ההודעה:

$$\Pr[plaintext = p|c] = \Pr[plaintext = p]$$

כלומר ידיעת ה-ciphertext לא תורמת דבר לפענוח ההודעה (הסיכוי שנמצא אותו שווה לסיכוי של plaintext כלשהו מהמרחב להיות ההודעה). לרוב התפלגות ההודעות ב-M אינו אחיד.

דוגמא: one-time pad:

- מרחב המפתחות הוא מרחב ההודעות, k (הצפנה סימטרית – מפתח הצפנה ופענוח) נבחר מתוכו באופן ראנדומלי. ההצפנה:

$$E_k(p) = c = p \oplus k, D_k(c) = c \oplus k = p$$

- כש-k נבחר אקראית מתוך המרחב, למעשה c מתפלג אחיד מעל M באופן ב"ת ב-p. ההצפנה מושלמת אך גודל המפתח כגודל ההודעה – גדול מאוד.

Theorem (Claud Shannon): אם מערכת הצפנה היא perfect cipher, גודל מרחב המפתחות כגודל מרחב ההודעות.

דוגמא: Vigenere cipher:

- ההצפנה: שמים את המפתח בחזרות רציפות תחת ההודעה, והמספר שמייצג כל תו במפתח (2 ל-b וכן הלאה) מסמל הזזה של האות המתאימה במיקום זה בהודעה המקורית.
- עבור גודל מפתח 1: אלגוריתם ההזזה; עבור גודל מפתח כלשהו l קטן מספיק מגודל ההודעה: ניתן לחלק את ההודעה ל-l הודעות (כל אחת מורכבת מקפיצות תוים במרחק l) ולנסות לפענח כמו פענוח הזזה לפי סטט' אותיות בודדות. אם גודל המפתח $\geq$ גודל ההודעה: זהו one-time pad.

רקע מתמטי 1:

סימונים:

- $a \equiv b (\mathrm{mod}\, m)$: $a - b$ מחלק את m.
- $a \bmod b$ (בלי סוגריים): שארית חלוקת a ב-b (בין 0 ל-$b-1$).

החוג $\mathbb{Z}_m$ (ring):

פעולות אריתמטיות מודולו m. פורמלית מיוצג: $(\mathbb{Z}_m, +, \cdot)$ כאשר $+, \cdot$ הן פעולות כפל וחיבור מודולו m, והאיברים הם $\{0, \dots, m-1\}$. תכונות:

- סגירות תחת חיבור וכפל: $a, b \in \mathbb{Z}_m \Rightarrow a+b \in \mathbb{Z}_m, \quad a \cdot b \in \mathbb{Z}_m$.
- קומוטטיביות ואסוציטיביות חיבור וכפל.
- דיסטריביוטיביות: $a \cdot (b+c) = a \cdot b + a \cdot c$.
- איבר נטרלי לחיבור: 0; איבר ניטרלי לכפל: 1.
- אם $a \in \mathbb{Z}_m$ אז $b \in \mathbb{Z}_m$ הוא הופכי כפלי שלו אם $a \cdot b = 1$. ל-0 אין הופכי כפלי, ול-1 יש: 1. לא לכל איבר בחוג יש הופכי כפלי.

טענה: אם $\gcd(a, m) = 1$ (המחלק המשותף המקסימלי של a,m הוא 1) אז ל-a יש הופכי כפלי ב-$\mathbb{Z}_m$.

חבורות, חוגים ושדות סופיים:

אקסיומות חבורה חילופית: *Commutative Groups*:

הגדרה: קבוצה G לא ריקה סופית/אינסופית של איברים ופעולה $+$ (סימון) על זוגות של איברים תהיה חבורה אם מקיימת:

- סגירות תחת $+$: $a + b \in G$
- אסוציטיביות: $(a+b)+c = a+(b+c)$
- קומוטטיביות: $a+b = b+a$
- קיים איבר ניטרלי 0: $a+0=a$
- לכל איבר קיים הופכי: $a + b_a = 0$

בחבורות לא קומוטטיביות או חבורות כפליות האיבר הניטרלי יסומן 1 והפעולה תסומן $\cdot$ (כמו כפל).

תתי חבורות:

$(H, +)$ תת חבורה של $(G, +)$ אם היא חבורה ו-$H \subseteq G$ (דוגמא: $(\mathbb{N}, +)$ אינה תת-חבורה של $(\mathbb{Z}, +)$ כי ב-$\mathbb{N}$ אין הופכי, $(\mathbb{Z}_{\text{even}}, +)$ כן).

טענה: אם $(G, +)$ חבורה סופית, $H \subseteq G$ ו-H סגורה לחיבור, אז $(H, +)$ חבורה בעצמה (חייבת להיות סופית, הדוגמא של קודם עם $\mathbb{N}, \mathbb{Z}$ מראה זאת).

משפט *Lagrange*: אם $(G, +)$ היא חבורה סופית ו-$(H, +)$ (עם אותה פעולה $+$) היא תת-חבורה שלה אז $|H| \mid |G|$ (גודל H מחלק את גודל G).

סדר: נסמן n חיבורים של a ב-a^n; נאמר כי a מסדר n אם $a^n = 0$ אבל לכל $m < n$ מתקיים $a^m \neq 0$.

טענה: בחבורה סופית, לכל a יש n שהוא לכל היותר סדר החבורה כך ש-$a^n = 0$ (הסדר של a הוא n).

הוכחה: אם נסתכל על הסדרה $a, a^2, a^3, \dots$, אז בסדרה זו חייבת להיות חזרה כיוון ש-G סופית, כלומר קיימים n_1, n_2 (שונים זה מזה) כך ש-$a^{n_1} = a^{n_2}$ (בה"כ $n_1 < n_2$), אז נקבל $0 = a^{n_1} - a^{n_1} = a^{n_2} - a^{n_1} = a^{n_2 - n_1} = a^l$, כאשר l הוא הסדר של a.

חבורה ציקלית:

נניח G היא חבורה סופית, ו-a איבר מסדר n, אזי $< a > := \{0, a, \dots, a^{n-1}\}$ היא תת חבורה של G. $< a >$ נקראת תת חבורה ציקלית שנוצרת ע"י a, ו-a הוא היוצר (*generator*) של החבורה. לפי משפט Lagrange, n מחלק את גודל G.

משפט *Fermat* הקטן:

עבור p ראשוני מתקיים לכל $a \in \{1, \dots, p-1\}$: $a^{p-1} \; mod\; p = 1$ (חזקה רגילה, לא חיבור $p-1$ פעמים). 0 אינו איבר בקבוצה.

הוכחה:

נתייחס לפעולת הכפל מודולו p ולקבוצה לעיל כ-$\mathbb{Z}_p^*$, שהיא חבורה כפלית.

בחבורה $\mathbb{Z}_p^*$ יש $p-1$ איברים, הסדר של כל $a \in \mathbb{Z}_p^*$, n, מחלק את $p-1$, כלומר $a^n \; mod\; p = 1$. לפי *Lagrange*: n מחלק את $p-1$, שהוא סדר החבורה. כלומר, קיים m טבעי כך ש-$n \cdot m = p-1$ (כפל טבעיים רגיל). מכאן: $(a^n)^m \; mod\; p = 1 \Rightarrow a^{p-1} \; mod\; p = 1$.

משפט: לכל p ראשוני החבורה $\mathbb{Z}_p^*$ היא חבורה ציקלית.

חוגים קומוטטיביים:

הקבוצה R היא חוג אם היא קבוצה לא ריקה עם שתי פעולות בינאריות: $+,\cdot$ המקיימת:

- סגירות תחת $+,\cdot$: $a,b \in R \Rightarrow a+b, a\cdot b \in R$
- אסוציאטיביות ביחס לכפל וחיבור: $(a+b)+c = a+(b+c), (a\cdot b)\cdot c = a\cdot(b\cdot c)$
- קומוטטיביות ביחס לכפל וחיבור: $a+b = b+a, a\cdot b = b\cdot a$
- קיים איבר ניטרלי ביחס לכפל וביחס לחיבור: $\forall a \in R: a+0 = a\cdot 1 = a$
- דיסטריביוטיביות: $a\cdot(b+c) = a\cdot b + a\cdot c$
- לכל איבר קיים הופכי ביחס לחיבור: $\forall a \exists b: a+b = 0$ (אין בהכרח הופכי כפלי)

שדות:

שדה הוא חוג עם יחידה בו לכל איבר שאינו 0 יש הופכי כפלי. למשל: $\mathbb{C}, \mathbb{R}, \mathbb{Z}_p$ כאשר p ראשוני. ההופכי הכפלי של a יסומן: a^{-1}.

<u>פולינומים מעל שדות</u>:

- לפולינום $f(x)$ מדרגה n יש מעל שדה $\mathbb{Z}_p$ לכל היותר n שורשים. משפט זה אינו נכון לסתם חוג עם יחידה (למשל ב-$\mathbb{Z}_{24}$ ל-$6x$ יש 6 שורשים).
- <u>שאריות פולינומים</u>: יהיו $f(x), g(x)$ פולינומים מדרגות n, m בהתאמה כך ש-$n \geq m$, אז קיים פולינום r(x) מדרגה <u>קטנה</u> מ-m ופולינום ייחודי נוסף h(x), שניהם מעל F, כך ש-$f(x) = h(x)\cdot g(x) + r(x)$. הפולינום r(x) מכונה <u>השארית</u> של f(x) מודולו g(x). השארית היא 0 אם $g(x)|f(x)$.

<u>שדה סופי</u>: שדה בו F (קבוצת האיברים) סופית. מתקיים: p ראשוני אמ"מ $\mathbb{Z}_p$ שדה סופי; לכל p ראשוני קיים שדה יחיד עם p איברים.

<u>מציין של שדה סופי (Characteristic)</u>: ה-n המינימלי הטבעי כך ש-$\underbrace{1+1+..+1}_{\text{n times}} = 0$. סימון: $\text{char}(F)$.

משפט: $\text{char}(F)$ הוא תמיד ראשוני.

שדות גלואה:

לכל חזקת מספר p ראשוני p^k ($k = 1,2,3 \dots$), קיים שדה סופי עם p^k איברים המקיים. סימון $F = GF(p^k)$. תכונות:

- $char(F) = p$
- $GF(p^k)$ ו-$\mathbb{Z}_{p^k}$ אינם אותו דבר!

<u>Symmetric Encryption: Stream & Block Ciphers</u>:

<u>Pseudo-Random Generators</u>:

<u>הגדרה</u>: פונ' פולי' $G: \{0,1\}^n \to \{0,1\}^m$ כאשר $m = n^c$ עבור $c > 1$ כלשהו המקיימת: הפלט של G אינו ניתן להבחנה ע"י מבחין פולי' מ-truly random string, כלומר המחרוזת המתקבלת אמורה להיראות אקראית כלפי מבחין מוגבל חישובית. ה-seed, n, הוא truly random seed, ולמעשה מרחיבים את האקראיות מ-n ל-m. PRG צריך להיות דטר' ופולי', כלומר הרגעין שלו ראנדומי אבל הפונ' G עצמה לא!

יהי D מבחין פולי' באורך הקלט m, מותר לו להטיל מטבעות והוא מכיר את פונ' ה-PRG. נאמר כי ה-PRG חזק אם:

$$\left| \Pr_{r\in\{0,1\}^n}[D(PRG(r)) = 1] - \Pr_{s\in\{0,1\}^m}[D(s) = 1] \right| < \frac{1}{n^c}$$

לכל $c > 1$. כלומר שמבחין מוגבל חישובית (פולי') לא יכול להבחין בין סתם מחרוזת ראנדומית באורך m לבין פלט של ה-PRG בהפרש גדול יותר מפונ' זניחה (1 חלקי פולינום כלשהו).

דוגמאות להצפנות:

- <u>Synchronous stream ciphers</u>: כל צד מחזיק ב-seed וב-PRG. השולח יוצר one-time-pad ע"י $PRG(seed)$, מצפין את ההודעה שלו ע"י XOR ושולח. המפענח יוצר גם הוא את $PRG(seed)$ ומפענח ע"י XOR עם ה-ciphertext שקיבל. חסרון: שני הצדדים צריכים להחזיק ב-seed, אם ביט יחיד הולך לאיבוד בתקשורת לא ניתן לפענח.
- <u>Asynchronous stream ciphers</u>: מתחילים מ-seed סודי ומייצרים ciphertext, ו-t הביטים האחרונים של ה-ciphertext נכנסים כקלט ל-PRG כהמשך למפתח הסודי. בשיטה זו מתגברים על אובדן ביטים בתקשורת: אחרי אובדן תקשורת, צריך לחכות שה-buffer יתמלא ב-t ביטים וניתן להמשיך לפענח. בסינכרונית המקבל צריך לדעת את מיקום הביטים שאבדו.

- LFSR: linear feedback shift registers: מערכ תהמייצרת מחרוזת ל-OTP, מאותחלת עם מפתח סודי c בעל L ביטים עבור L שלבים. המפתח המגונרט מקיים: $s_j = \oplus_{i=1}^{L} c_i s_{j-i}$ (סכום מודולו 2). שיטה זו מהירה אך לא בטוחה – עם מספיק ביטים ניתן לגלות את המפתח.

פונקציה חד כיוונית:

פונקציה שקל להצפין עמה: $x \to f(x)$ אך קשה לפענח אותה ללא מפתח: $f(x) \to x$.

Block Ciphers:

הצפנת בלוק קלט לבלוק פלט, כאשר גודלם לרוב זהה (גודל ה-ciphertext יהיה ≤ גודל ה-plaintext). גדלי הבלוקים בפועל הם בד"כ $n = 64$ (DES) או $n = 182$ (AES). שיטות הצפנת בלוקים:

- ECD: electronic code book: כל בלוק עובר דרך פונ' ההצפנה עם המפתח k: $P_i \to E_k \to C_i$. בעיה: חזרות של בלוקים יכולות לתת מידע ליריב כיוון שבלוקים זהים בהודעה נותנים בלוקים זהים בהצפנה. אין זו חולשת E_k אלא האופן בו משתמשים בה.
- CBC mode: cipher block chaining: מתחילים מ-S_o מילה קבועה (באורך הבלוק). האלג': $C_1 = E_k(P_1 \oplus S_0), \forall i > 1: C_i = E_k(P_i \oplus C_{i-1})$. מערכת זו אסינכרונית, כלומר בידיעת ciphertext של בלוק מסויים ניתן לפענח ממנו את השאר (בהינתן k). אם E היא פרמוטציה פסאודו-ראנדומית אז CBC עמידה בפני chosen plaintext attacks.
- OFB: output feedback mode: גם כאן מתחילים מ-s_0 מילה קבועה, כאשר לכל i: $C_i = P_i \oplus s_i,\ s_i = E_k(s_{i-1})$. כלומר ה-s מכל שלב מהווה מקור ל-s לשלב הבא. בדומה ל-CBC שיטה זו היא אסינכרונית ועמידה בפני ההתקפה הנ"ל.

עקרונות תכנון ל-Block Ciphers:

נסמן "פירוק" של E_k לתתי פונ' f_{k_i} כאשר k_i חלק מהמפתח k. בהינתן המפתח k חישוב הצפנה ופענוח הוא קל (מהיר). ללא המפתח k E_k אמורה להיראות כמו תמורה אקראית על מרחב ההודעות. נרצה להתמודד מול יריב המסוגל:

- לצפות בזוגות P_i, C_i.
- Chosen plaintext attacks: מותר ליריב לבחור מספר סביר של זוגות P_i, C_i.

עם יכולות אלו נרצה שהיריב לא יוכל לקבל אינפורמציה על k ואפילו לא לדעת על זוג שלא ראה P_j, C_j האם השני הוא תוצר הפעלת E_k על הראשון. זוהי תמורה פסאודו-אקראית.

DES: Data Encryption Standard:

ההצפנה מורכבת מאיטרציות (סיבובים), כאשר f_{k_i} היא פונ' ההצפנה של הסיבוב ה-i. מחלקים תחילה את ה-Plaintext לשני חלקים $L_0 R_0$. האלג':

- $L_i = R_{i-1}$
- $R_i = L_{i-1} \oplus f_{k_i}(R_{i-1})$

אם f היא פונ' פסאודו אקראית אזי לאחר 4 הרכבות (הנקראות הרכבות Feistel) המנגנון שואף לתמורה פסאודו אקראית. שחזור:

- $R_i = L_{i+1}$
- $L_i = R_{i+1} \oplus f_{k_i}(L_{i+1})$

AES: Advanced Encryption Standard:

הצפנת בלוקים סימטרית עם מפתחות באורכים: 128,192,256 ביטים העמידה בפני כל ההתקפות הידועות. בה"כ נסתכל על שיטת 128 ביטים. כל בלוק בגודל 128 ביט מוצפן עם מפתח באורך 128 ביט ב-10 שלבים. כל מצב הוא 128 ביטים המוחזקים במטריצה 4×4 של בתים, שכל אחד מהאלמנטים בה הוא איבר ב-$GF(2^8)$. בכל סיבוב מ-10 הסיבובים:

- Substitution: מחליפים את כל איברי המטריצה בהופכיים שלהם, 0 נשאר עצמו. זו פעולה חח"ע ולא לינארית.
- הזזת שורות: בשורה ה-i מבצעים הזזה ציקלית של i מקומות.
- ערבוב עמודות ע"י פעולות אריתמטיות כלשהן.
- XOR round key: מבצעים XOR על המפתח של הסיבוב הנוכחי עם ה-state כדי לקבל state חדש.

רקע מתמטי 2:

אלגוריתם GCD של אאוקליד:

בהינתן r_0, r_1 טבעיים, האלגוריתם: מגדירים סדרת פעולות מהצורה $r_{i+2} = r_i \bmod r_{i+1}$ על i החל מ-0, ועוצרים כאשר $r_i = 0$ ו-$gcd = r_{i-1}$.

סיבוכיות: נניח $|r_0| = n$ (n ביטים). $r_2 = r_0 \bmod r_1$ ולכן קיים $c \geq 1$ טבעי כך ש: $r_0 = c \cdot r_1 + r_2$ כאשר $r_1 > r_2$. לכן $r_0 > 2r_2$, ומכאן $r_2 \leq \frac{r_0}{2}$, $r_3 \leq \frac{r_1}{2}$, $r_4 \leq \frac{r_0}{4}$, $r_6 \leq \frac{r_0}{8}$ וכן הלאה: $r_{2i} \leq \frac{r_0}{2^i}$ - כל שני סיבובים קטן בפקטור 2. מכאן מספר הסיבובים לכל היותר: $2n$.

אלגוריתם אאוקליד המוכלל:

אם r_o ו-r_1 טבעיים כך ש-$\gcd(r_0, r_1) = g$ אז יש x, y שלמים כך ש-$r_0x + r_1y = g$. הוכחה לכך היא פשוטה (באינדוקציה). אם a, m זרים אז $\gcd(a, m) = 1$ ואז יש x, y כך ש-$ax + my = 1$ (אלג' אאוקלידי המוכלל מוצא אותם) ומתקיים ש-x הוא ההופכי הכפלי של a מודולו m.

הפונקציה $\phi(m)$: פונקציית totient של אויילר, מסמנת את מספר המספרים ב-$[1, \dots, m]$ הזרים ל-m. מתקיים:

- אם k, l זרים אז $\phi(l \cdot k) = \phi(l) \cdot \phi(k)$.
- אם p ראשוני אז $\phi(p) = p - 1$.
- אם $m = p^l$ ו-p ראשוני אז מספר הלא זרים הוא p^{l-1} (על כל p אחד יש זר אחד) ולכן $\phi(m) = p^l\left(1 - \frac{1}{p}\right)$.
- אם $m = p_1^{\alpha_1} \cdot \dots \cdot p_k^{\alpha_k}$, אז $\alpha_i \geq 1$, $\varphi(m) = \prod_{i=1}^{k}\left(p_i^{\alpha_i} - p_i^{\alpha_i - 1}\right)$.

בהינתן פירוק, חישוב $\phi(m)$ קל, אך ללא פירוק חישוב $\phi(m)$ קשה כמו פירוק m. $\phi(m)$ הוא מספר האיברים בחוג הכפלי $\mathbb{Z}_m^*$ (אוסף כל האיברים שיש להם הופכי כפלי מודולו m) כלומר סדר החבורה.

Iterated ciphers, Message Authentication codes, Crypt. Hash Func.:

חיזוקים להצפנת Block Cipher נתונה:

Iterated Ciphers:

הצפנת הבלוק פעמיים עם שני מפתחות שונים: $P_i \to E_{k_1} \to E_{k_2} \to C_i$. מרחב המפתחות גדל למפתחות עם 2n ביטים, כלומר זמן פיצוח נאיבי הוא $O(2^{2n})$. בעיה: אם פונ' ההצפנה סגורה להרכבה, למשל XOR: $p \oplus k_1 \oplus k_2$ שקול ל-$p \oplus k_3$ עבור $k_3 = k_1 \oplus k_2$ ($|k_3| = |k_1 \oplus k_2|$).

Meet in the Middle attack:

יהיו x, y זוג ptext ו-cptext כך ש-$y = E_{k_2} \circ E_{k_1}(x)$. מראש ישנם $2^n \cdot 2^n$ זוגות פוטנציאלים ל-(k_1, k_2), והתקפה זו מצמצמת את מס' הזוגות האפשריים לבערך 2^n. שימוש בזוגות נוספים יעשה כדי לצמצם את הרשימה לזוג יחיד. האלג':

- לכל k_1 אפשרי מחשבים: $z_{k_1} = E_{k_1}(x)$ ושמים את (z_{k_1}, k_1) ברשימה L_1 ממויינת לפי z_{k_1}. זמן וזיכרון $O(n \cdot 2^n)$.
- לכל k_2 אפשרי מחשבים את $t_{k_2} = D_{k_2}(y)$ ושמים את (t_{k_2}, k_2) ברשימה L_2 ממויינת לפי t_{k_2}. אותה סיבוכיות.
- עבור הזוג הנכון מתקיים: $z_{k_1} = t_{k_2}$. יהיו יותר מזוג מפתחות אחד שיקיים זאת, ולכן נשתמש ב-x, y נוספים עד שנגיע לזוג הנכון. מציאת הזוגות ברשימות (ממויינות) היא $O(2^n)$.

היוריסטיקה: ניתן להניח כי E אינה תמורה אקראית. ההסתברות ש-$E_{k_1}(x) = D_{k_2}(y)$ היא $\frac{1}{2^n}$. עבור מפתח באורך l, מספר זוגות המפתחות הוא 2^l, ואז תוחלת מספר הזוגות המקיימים את השוויון הוא 2^{2l-n}, ואם $n = l$ התוחלת תהיה 2^n. עבור 2 זוגות (x_i, y_i) הסיכוי לשניהם הוא $\frac{1}{2^n} \cdot \frac{1}{2^n} = \frac{1}{2^{2n}}$, ותוחלת מספר ההתנגשויות תקטן עוד יותר ל-2^{2l-2n} - זוגות בודדים של מפתחות.

Triple Cipher:

- האלג': $C_i = E_{k_3} \circ D_{k_2} \circ E_{k_1}(P_i)$: השימוש ב-D באמצע הוא לשם תאימות לאחור, כך שניתן יהיה לשלוח single-cipher (שימוש במפתח יחיד k בהצפנה זו: $E_k \circ D_k \circ E_k(x) = E_k(x)$).
- גם כאן ניתן לנסות לבצע meet-in-the-middle אך מצד אחד יהיו שני מפתחות מצד שני אחד, והסיבוכיות תהיה בגודל מפתח 2n.

From Encryption to Authentication:

תרחיש: אליס רוצה לשלוח הודעה לבוב. פראן (Forger) עלולה להתחזות לאליס ולשלוח הודעות לבוב בתור אליס. בוב רוצה להבחין בכך.

הערה: סודיות אינה מבטיחה אימות ולהיפך.

סימונים:

- A: אלגוריתם אימות, מסומן MAC, V: אלגוריתם וידוא, k מפתח ו-M מרחב ההודעות.
- הודעה היא זוג: $\left(m, A_k(m)\right)$ (m הודעה לא מוצפנת), כאשר $A_k(m)$, המסומן גם $MAC_k(m)$, הוא ה-authentication tag של m.

<u>דרישות מערכת אימות</u>:

- קונסיסטנטיות: $V_k \circ A_k(m) = accept$.
- יריב המוגבל פולי' (למשל) לא יוכל לבנות זוג מתאים $\left(m, MAC_k(m)\right)$ אלא בהסתברות זניחה, גם לאחר שראה n זוגות אמיתיים (יודע את אלג' ה-MAC אך לא את k).

<u>שימושי MAC</u>:

CBC-MAC: Cipher Block Chaining:

- עבור הודעה המחולקת לבלוקים m_i מייצרים את c_i (עם $seed = 00 \dots 0$). זורקים את כל התוצרים פרט לאחרון c_n שהוא $MAC_k(m)$.
- <u>משפט</u>: אם E_k פסאודו-אקראית, שיטה זו טובה אם מספר הבלוקים המרכיבים את ההודעה קבוע וידוע מראש, לא בטוח אחרת. דוגמא לזיוף:

  אם נשלח תחילה (m_1, c_1) ולאחר מכן (c_1, c_2) אז ניתן לשלוח את ההודעה $\left(m_1 \circ \vec{0}\,, c_2\right)$ (∘ - שרשור), כי: $E_k(\underbrace{c_1 \oplus 0}_{=c_1}) = c_2$.

שיטות להתגבר על כך:

- שרשור מספר הבלוקים לתחילת ההודעה, כלומר השמת MAC על $(n, m_1, \dots, m_n)$. חסרון: אורך ההודעה צריך להיות ידוע מראש.
- שימוש ב-k_2 (מפתח נוסף) וחישוב: $MAC_{k_1,k_2}(m) = E_{k_2} \circ MAC_{k_1}(m)$ (מומלץ).

Cryptographic Hash Functions:

פונקציות הממפות תחום גדול לטווח קטן יותר. פונקציות אלו אינן חח"ע (כמובן) ו<u>אין מפתח סודי</u>. דרישות מפונ' hash קריפטוגרפית:

- לכל y קשה למצוא x כך ש-$h(x) = y$.
- Weak collision resistance: לכל x_1 קשה למצוא $x_2 (\neq x_1)$ כך ש-$h(x_1) = h(x_2)$ (collision).
- Strong collision resistance: קשה למצוא זוג x_1, x_2 כך ש-$h(x_1) = h(x_2)$.

פונ' hash קריפטו' $h: \{0,1\}^n \to \{0,1\}^m$ הינה בעלת מעט התנגשויות, מהירה לחישוב ובד"כ $n = 512, m \geq 160$.

<u>פרדוקס יום ההולדת</u>: ההסתברות שמתוך קבוצה של 23 איש לשניים מהם יש אותו יום ההולדת היא גדולה מ-$\frac{1}{2}$.

<u>Quadratic Residues, The Discrete Logarithm Problem</u>:

מציאת איבר פרימיטיבי:

בכל שדה $GF^*(p^k)$ קיים איבר פרימיטיבי. מציאת איבר פרימיטיבי:

- מגרילים איבר בשדה שאינו 0, יש לו סיכוי גבוה להיות פרימיטיבי.
- בודקים את הסדר שלו: נניח נתון הפירוק $\Pi_{i=1}^{s} p_i^{e_i} = p^k - 1$, כאשר p_i ראשוניים ו-$e_i \geq 1$ (בעייתיות: פירוק היא בעיה קשה, לא תמיד נתון).
- מסמנים את סדר האיבר שהגרלנו x: $|x| = m = \Pi_{j=1}^{s} p_j^{f_j}$. בודקים האם קיים j כך ש-$f_j < e_j$ כך ש: $x^{\frac{p^k-1}{p_j}} = 1$.
- אם כן – זה אומר שסדר האיבר קטן מסדר החבורה, אז האיבר לא פרימיטיבי. אחרת סדר האיבר הוא סדר החבורה – האיבר כן פרימיטיבי.

<u>שאריות ריבועיות – Quadratic Residues</u>:

<u>הגדרה</u>: $m \geq 2, x \neq 0$ נאמר כי $x \in \mathbb{Z}_m$ הוא שארית ריבועית (מודולו m) אם קיים $y \in \mathbb{Z}_m$ כך ש-$y^2 \equiv x (\mathrm{mod}\ m)$.

<u>משפט אויילר</u>: $x \in \mathbb{Z}_p^*$ הוא שארית ריבועית $\Leftrightarrow x^{\frac{p-1}{2}} \equiv 1 (\mathrm{mod}\ p)$.

<u>טענות</u>: עבור $p > 2$ ראשוני, ב-$\mathbb{Z}_p^*$ יש $\frac{p-1}{2}$ שאריות ריבועיות, ואם g יוצר אז הן $\{g^0, g^2, g^4, \dots, g^{2i}, \dots, g^{p-3}\}$.

<u>אלגוריתם יעיל לבדיקת QR ב-$\mathbb{Z}_p^*$</u>: מחשבים את $x^{\frac{p-1}{2}}$ ע"י העלאות חוזרות ונישנות בחזקה של $x \to x^2 \to x^4 \to \cdots$. האלג' מקבל כקלט את x, p ולכן אורכו הוא $2 \log p$. סה"כ נזדקק ל-$\log p$ העלאות בחזקה, שכל אחת עולה $\log^2 p$, לכן סה"כ הסיבוכיות היא $O(\log^3 p)$.

<u>אלגוריתם יעיל לבדיקת QR ב-$\mathbb{Z}_m^*$, $m = pq$ לא ראשוני</u>: x הוא ב-QR אמ"מ x הוא ב-QR גם ב-$\mathbb{Z}_p^*$ וגם ב-$\mathbb{Z}_q^*$. בעיה זו פתירה באופן יעיל אם הפירוק של m נתון, אך אם הוא לא נתון זו בעיה קשה.

The Discrete-Log problem:

הגדרה: תהי $G =< g \geq \{g^1, \dots, g^{|G|}\}$ חבורה ציקלית; אם $x \in G$ אז קיים i מינימלי כך ש-$g^i = x$, ואז: $Dl_g(x) \coloneqq i$ – ומציאת הלוגריתם הדיסקרטי היא בעיה קשה (דיסקרטי – משום שזה כפל בחבורה). זו דוגמא ל**פונקציה חד כיוונית**: $g^i \to i$ קל, $i \to g^i$ קשה.

Public Key Cryptography:

שימוש בשני מפתחות: מפתח הצפנה פומבי, מפתח פענוח סודי.

Diffie Hellman key-exchange:

עובדים מעל חבורה $\mathbb{Z}_p^*$ עם p ראשוני גדול מהצורה: $p = 1 + (small\ factors) \cdot (\boldsymbol{one}\ large\ factor)$, ועם g יוצר בחבורה. הפרוטוקול:

- אליס מחשבת את $x = g^a (\mathrm{mod}\, p)$, $a \in [0, \dots, p-2]$ מפתח סודי של אליס, ושולחת לבוב.
- בוב מחשב $y = g^b (\mathrm{mod}\, p)$, $b \in [0, \dots, p-2]$ מפתח סודי של בוב, ושולח לאליס.
- אליס מחשבת: $x^b = (g^a)^b = g^{ab}$, בוב מחשב: $y^a = (g^b)^a = g^{ab}$ המפתח המשותף.

הערות:

- בטיחות DH key-exchange היא לכל היותר כמו קושי DL ב-$\mathbb{Z}_p^*$ (בהינתן פתרון יעיל ל-DL ניתן לפרוץ ביעילות את DH).
- מיפוי $a \to g^a (\mathrm{mod}\, p), b \to g^b (\mathrm{mod}\, p)$ הינו יחיד (g יוצר).
- אינפורמציה שניתן ללמוד על המפתח הפרטי מהפומבי: g^a הוא ב-QR $\Leftrightarrow a$ הוא זוגי (ה-lsb שלו היא 0). כך ניתן ללמוד גם על b ולמעשה על ab.

מספרים ראשוניים ובדיקת ראשוניות:

משפט: עבור $m = pq$ כאשר p, q ראשוניים, ב-$\mathbb{Z}_m^*$ אין איברים פרימיטיביים.

משפט המספרים הראשוניים:

- ישנם $\aleph_0$ ראשוניים (הוכחה: אם יש רק n ראשוניים תמיד נוכל ליצור מספר $p_1 \cdot \dots \cdot p_n + 1$ שהוא ראשוני / גורמיו אינם $p_1, \dots, p_n$).
- צפיפות הראשוניים: יהי $\pi(x)$ מספר הראשוניים עד x, אז מתקיים: $\pi(x) \cong \frac{x}{\ln x}$.

  יהי p_n הראשוני ה-n-י, אז מתקיים: $n \ln n + n(\ln \ln n - 1) < p_n < n \ln n + n \ln \ln n$.

  $\Leftarrow$ אם בוחרים מספר בן n ספרות באקראי, סיכויו להיות ראשוני הם בערך $1/n$.
- אנלוגיה לפולינומים ב-$GF(p^k)$: מס' הפולינומים האי-פריקים מתוך כלל p^k הפולינומים הוא בערך $\frac{p^k}{k}$.

בדיקת ראשוניות:

המשפט הקטן של פרמה: אם p ראשוני ו-$1 \leq a \leq p-1$ אז $a^{p-1} \equiv 1 (\mathrm{mod}\, p)$; מכאן: אם קיים $a \in \mathbb{Z}_m$ כך ש-$a^{m-1} \neq 1 (\mathrm{mod}\, m)$ אז m אינו ראשוני. בעיה: מספרי קרמייקל מקלקלים תיאוריה זו.

מספרי קרמייקל: מספרים מהצורה $m = p_1 \cdot \dots \cdot p_k$ ($k \geq 3$) כך שלכל i: $p_i - 1 \mid m - 1$ - מספרים אלו מכשילים את מבחן פרמה. לכן, בדיקת ראשוניות לפי פרמה צריכה לבדוק שהמספר אינו מספר קרמייקל.

מבחן מורחב לפריקות: מבצעים 3 בדיקות על $2 \leq a \leq m-1$:

1) אם $\gcd(m, a) > 1$ אז m פריק.
2) אם $a^{m-1} \neq 1 (\mathrm{mod}\, m)$ אז m פריק (מבחן פרמה).
3) $a^2 \equiv 1 (\mathrm{mod}\, m)$ וגם $a \neq m-1$ אז m הוא פריק.

שיפור הבדיקה: נניח כי $m - 1 = 2^k \cdot r$ (r אי זוגי):

- בוחרים $a = b^r$ ומתקיים: $\underbrace{(\dots (b^r)^2 \dots)^2}_{k\ times} = b^{m-1}$ – מעלים בריבוע את b^r כ-k פעמים, ואם בסוף הבדיקה $b^{m-1} \neq 1 (\mathrm{mod}\, m)$ זהו עד לכך ש-m פריק (לפי מבחן 2).
- אם לא, מגדירים: $\begin{cases} a_0 = b^r \\ a_i = a_{i-1}^2 \end{cases}$ וכך $a_k = b^{m-1} (\mathrm{mod}\, m) = 1$ (כי מבחן 2 לא התקיים), ונסמן j האינדקס הקטן ביותר בו: $a_j \equiv 1 (\mathrm{mod}\, m)$.

  אם $j > 0$, $a_{j-1} \neq -1 (\mathrm{mod}\, m)$ אז a_{j-1} הוא עד מסוג 3, ו-m הוא פריק.

b המקיים את אחד משני הנ"ל הוא עד חכם.

משפט Rabin: אם m פריק, לפחות $3/4$ מהמספרים בטווח $1, \dots, m$ הם עדים חכמים.

מבחן Miller-Rabin: משתמש במשפט רבין כדי לבדוק פריקות בסבירות גבוהה. יהי m מספר בן n ביטים, נבצע 100 פעמים: בוחרים באקראי $1 < b < m$ ובודקים האם הוא עד חכם. אם אחד או יותר מה-b הוא עד חכם, מחזירים ש-m פריק, אחרת שהוא ראשוני. עבור m ראשוני, תמיד יוחזר שהוא ראשוני. עבור m פריק, בסיכוי לטעות קטן מ-$\left(\frac{1}{4}\right)^{100}$ יוחזר שהוא ראשוני. הבדיקה פולי' באורך הקלט. אלג' ב-RP (הסתברותי עם טעות ח"צ).

קיים אלגוריתם דטר' לבדיקת ראשונית (ומכאן פריקות): כמה הודים המציאו אותו, אבל בפרקטיקה עדיין משתמשים במילר-רבין.

הכפלת שלמים ופקטוריזציה כפונקציה חד-כיוונית:

פונקציה חד כיוונית היא פונ' שקל לחשב אותה בכיוון אחד, אך קשה לחשב את ההופכית שלה. קל לבחור שני ראשוניים ולחשב את מכפלתם $m = pq$, אך הפירוק $m \to p, q$ קשה. RSA מתבסס על כך:

RSA:

- האינפורמציה הפרטית: של בוב היא p, q - שני ראשוניים גדולים אקראיים.
- האינפורמציה הפומבית: $m = pq$, חזקה e שזרה ל-$\phi(m) = (p-1)(q-1)$ - מספר האיברים ב-$\mathbb{Z}_m^*$.
- עוד אינפורמציה פרטית: d הזר ל-$\phi(m)$ כך ש-$ed \equiv 1 (\mathrm{mod}\ \phi(m))$ - ניתן לחשב את d באמצעות $xgcd$.
- הצפנה עבור הודעה $A \in \mathbb{Z}_m$: $C := A^e (\mathrm{mod}\ m)$ - אליס מצפינה לבוב הודעה שלה עם המפתח הפומבי שפרסם.
- פענוח C: בוב מפענח את ההודעה של אליס ע"י חישוב: $C^d = A^{ed} (\mathrm{mod}\ m) = A^{ed (\mathrm{mod}\ \phi(m))} = A^1$.

הערה: עבור $A \in \mathbb{Z}_m \backslash \mathbb{Z}_m^*$ זה לא מתקיים, אך לא הסבירות להתקל ב-A כזה נמוכה.

הערה: בהינתן $e, m = pq$ (המידע הפומבי), חישוב $\phi(m)$ ומציאת d שקולה לפירוק.

משפט השאריות הסיני:

יהיו:

- $m_1, \dots, m_k$ טבעיים זרים בזוגות.
- $a_1, \dots, a_k$ טבעיים כך ש-$0 \le a_i \le m_i - 1$.

אזי יש x טבעי כך שלכל i מתקיים $x \equiv a_i (\mathrm{mod}\ m_i)$ ובתחום $\left[0, \Pi_{i=1}^k m_i - 1\right]$ הוא יחיד.

שורשים ריבועיים של 1 ב-$\mathbb{Z}_m^*$:

ב-$\mathbb{Z}_p^*$ (וב-$\mathbb{Z}_q^*$) יש שני שורשים ריבועיים טרוויאלים ל-1: $1, -1 = p-1$ (בהתאמה עם q). לפי משפט השאריות הסיני: ב-$\mathbb{Z}_m^*$ ל-1 יש 4 שורשים ריבועיים. באופן כללי, אם $z \in \mathbb{Z}_m^*$ ריבוע, אז קיים $t \in \mathbb{Z}_m^*$ כך ש-$t^2 \equiv z (\mathrm{mod}\ pq)$ ומכאן ישנם 4 שורשים ריבועיים:

- $t \cdot (q-1)$
- $t \cdot (p-1)$
- $t \cdot (-1) = t(pq-1)$
- $t \cdot 1 = t$

ומכאן: ההעתקה $x \to x^2 (\mathrm{mod}\ pq)$ היא 4 ל-1, כלומר $\frac{1}{4}$ מהאיברים הם ריבועים ו-$\frac{3}{4}$ הם לא (עבור $\mathbb{Z}_{pqr}^*$ נקבל העתקה 8 ל-1, p, q, r ראשוניים).

חזרה ל-RSA:

טענה: אם e זר ל-$\phi(m) = (p-1)(q-1)$ אז ההעתקה $x \to x^e (\mathrm{mod}\ m)$ היא חח"ע ועל $\mathbb{Z}_m^*$ - ואז יכולה לשמש בסיס להצפנה.

הערות:

- חישוב $x^{ed} (\mathrm{mod}\ pq)$ שקול לחישוב $x^{ed (\mathrm{mod}\ (p-1)(q-1))} (\mathrm{mod}\ pq)$.
- RSA הוא דטרמיניסטי, ולכן ניתן להכניס padding אקראי כדי לבלבל את היריב.
- RSA סגור תחת כפליות: $E(P_1 \cdot P_2) = E(P_1) \cdot E(P_2)$ כי $(xy)^e (\mathrm{mod}\ m) = x^e \cdot y^e (\mathrm{mod}\ m)$ - פגיעות ל-chosen ciphertext attacks. גם כאן ניתן לפתור זאת ע"י padding ראנדומי.

חתימות:

רעיון החתימה הוא צירוף מחרוזת דיגיטלית כלשהי להודעה המהווה אותנטיקציה שההודעה אכן נשלחה ממי שנטען שהיא נשלחה ממנו.

פתרון DH:

- תהי E פונ' הצפנה פומבית ו-D פונ' פענוח פרטית. החותם שולח את הזוג $\big(M, D(M)\big)$.
- המקבל משתמש בפונ' הפומבית כדי לבדוק לראות שאכן $E\big(D(M)\big) = M$.
- כיוון ש-D היא פונ' פרטית של החותם, הוא היחיד שיכול לחתום עמה.

הנ"ל קל לזיוף.

<u>פתרון ע"י שימוש ב-Hash</u>: עבור H פונ' hash עמידה בפני התנגשויות, ניתן לחתום כך: $\left(M, D\left(H(M)\right)\right)$. האימות מתבצע ע"י החישוב כמו קודם, ובדיקה שמה שמתקבל שווה ל-$H(M)$.

<u>סכמת חתימה כללית</u>:

- גנרציה של מפתחות: שלב זה חייב להשתמש ב-truly random bits.
- אלגוריתם חתימה A
- אלגוריתם וידוא V המחזיר accept/reject

<u>אלגוריתמים לפירוק</u>:

אלגוריתם ρ של Pollard:

- סיבוכיות: $2^{\frac{n}{4}}$ (קיימים טובים יותר).
- תהי F פונקציה ו-$x \in \mathbb{Z}_m$, נסתכל על הסדרה: $x, F(x), F^2(x), \dots$ - כיוון ש-$\mathbb{Z}_m$ מרחב סופי, המסלול חייב לחזור על עצמו (עד הכניסה ללולאה זה ה"זנב" של הסדרה, ומשם זו הלולאה).
- אם F אקראית, מפרדוקס היומולדת נקבל כי אורך הזנב ואורך הלולאה שניהם בערך $\sqrt{\frac{\pi}{8}m}$ (תוחלת על F ועל כל x).
- תהיינה $F_p = F(\bmod p), F_q = F(\bmod q)$ ב"ת. אם סוגרים לולאה ב-F_p אך לא ב-F_q זה אומר כי קיימים y, z כך ש-$y = z(\bmod p)$ וגם $y \neq z(\bmod q)$, ולכן $y - z$ מתחלק ב-p אך לא ב-q.
- גילוי התנגשות ע"י הרצת שני מצביעים, שקצב הראשון צעד אחד וקצב השני שני צעדים – שניהם יתנגשו באופן ודאי.

<u>Elgammal public-key cryptosystem</u>:

הצפנת Elgammal מבוססת על DL אך לא ידוע ששקולה אליה (כמו היחס בין RSA ופירוק).

<u>אלגוריתם Elgammal PKC</u>:

- יהי p גורם ראשוני גדול שפירוקו ידוע, עדיף מהצורה $p = 2q + 1$ כאשר q הוא ראשוני ויהי g איבר פרימיטיבי ב-$\mathbb{Z}_p^*$. p, g פומביים.
- בוב בוחר באקראי $a \in [0, \dots, p-2]$ ומפרסם את $\beta = g^a$, a המפתח בפומבי שלו.
- לאליס הודעה m, והיא בוחרת באקראי $k \in [0, \dots, p-2]$ ומחשבת את: $g^k(\bmod p)$, $m \cdot \beta^k(\bmod p)$ ושולחת: $(g^k, m\beta^k)$.
- בוב מפענח כך: $(g^k)^a = (g^a)^k = \beta^k(\bmod p)$. ע"י $xgcd$ בוב יכול לחשב את β^{-k}, גם מבלי לדעת את k. כעת הוא יכול לחשב את m.

דרך לפרוץ: מציאת k מתוך g^k, אך זו בעית DL. תכונות:

- ההצפנה היא אקראית, ואם אליס תשתמש באותו k פעמיים ניתן יהיה לחשב את היחס בין 2 שתי ההודעות: $\begin{cases} g^k, m_1\beta^k \\ g^k m_2 \beta^k \end{cases} \Rightarrow \frac{m_1}{m_2}$.
- אי עמידות ל-chosen ciphertext attack: עבור מציאת m מתוך $(g^k, m\beta^k)$, תוקף יכול לבקש את המקור של $(g^k, sm\beta^k)$ וכך לגלות את sm ומשם בקלות לגלות את m.
- המערכת היא מולטיפליקטיבית.

<u>דליפת אינפורמציה חלקית ב-Elgammal</u>: כיוון ש-$g^a = \beta, g^k$ הם אינפורמציה פומבית, ניתן (כמו ב-DH) לגלות את ה-lsb של a ושל k. מכאן ניתן ללמוד האם β^k ב-QR והאם $m\beta^k$ ב-QR, וכיוון שהחבורה כפלית ניתן ללמוד האם m ב-QR ולמעשה את ה-lsb של m. לכן כדאי לקחת תת חבורה של כל מי שב-QR (מחצית מהאיברים).

<u>Elgammal signature scheme</u>:

<u>יצירת המפתחות</u>:

יהי ראשוני גדול $p = 2q + 1$ כאשר q גם ראשוני גדול (p בגודל 1024 ביטים) ותהי H פונ' hash עמידה בפני התנגשויות.

- בוב בוחר איבר פרימיטיבי $g \in \mathbb{Z}_p^*$, ובוחר באקראי $x \in [0 \dots p-2]$.
- לאחר מכן מחשב $y = g^x(mod\, p)$, והוא חלק מהמפתח הפומבי.

<u>אלג' החתימה</u>:

תהי M הודעה ותהי $m \coloneqq H(M)$. בוב בוחר באקראי $k \in [0 \dots p-2]$ **זר** ל-$p-1$ (מתוך $\phi(p-1) \geq q-1$ כאלה).

- בוב מחשב את $r = g^k (mod\ p)$
- בוב מחשב את $s = (m - rx) \cdot k^{-1}\ mod\ p-1$
- בוב מוציא את r, s.
- החתימה היא: M, r, s.

<u>בדיקה של אליס שהחתימה היא אכן של בוב</u>:

- מפתח פומבי: $\mathrm{p, g, y = g^x}$ ידועים.
- מפתח פרטי: x.
- אליס מקבלת M, s, r.

אליס בודקת האם $0 < r < p$ וגם $\mathrm{y^r r^s = g^m (mod\ p)}$. אם כן, אליס מקבלת, אחרת היא דוחה. הסבר:

- כיוון ש: $\mathrm{s = (m - rx)k^{-1}\ mod\ p - 1}$ מתקיים: $\mathrm{sk + rx = m\ mod\ p - 1}$
- כיוון ש: $\mathrm{r = g^k}$ מתקיים: $\mathrm{r^s = g^{ks}}$
- כיוון ש: $\mathrm{y = g^x}$ מתקיים: $\mathrm{y^r = g^{rx}}$

$\Leftarrow \mathrm{y^r \cdot r^s = g^{rx} \cdot g^{ks} = g^{sk+rx} = g^m}$

<u>חלוקת סוד: Secret Sharing</u>:

$n-out-of-n$ secret sharing:

נניח נתון סוד חשוב s ו-trusted dealer המחזיק בו. רוצים לחלק את הסוד בין n משתתפים כך ש:

- רק יחד יוכלו כולם לשחזר את הסוד.
- אף תת קבוצה של $n-1$ לא יוכלו לשחזר את הסוד או להסיק מחלקה משהו על s.

לדילר פונקציה $F(S, r) \to (p_1, \dots, p_n)$ שבהינתן סוד וחלק אקראיות מייצר חלקים כך שהמשתתף ה-i מקבל את p_i. דרישות הסודיות:

$$\forall x. \forall i. Pr[S = x] = Pr[S = x | \{p_1, \dots, p_n\} \backslash \{p_i\}]$$

<u>פתרון אפשרי</u>: הדילר בוחר $n-1$ ביטים באקראי ובאופן ב"ת $p_1, \dots, p_{n-1}$ והביט האחרון הוא: $p_n = s \oplus \sum_{i=1}^{n-1} p_i (\mathrm{mod}\ 2) = s \oplus_{i=1}^{n-1} p_i$.

- שחזור: $s = \oplus_{i=1}^{n} p_i$
- חוסר אינפ' מתת קבוצה: כל $n-1$ ביטים מתפלגים אחיד באופן ב"ת ללא קשר ל-s ולכן ההסתברות למצוא את s נשארת זהה.

באופן כללי:

- יהי U מרחב הסודות, $|U| = m$, בה"כ $\mathbb{Z}_m = \{0, \dots, m-1\}$. יהי סוד $S \in U$.
- הדילר בוחר באקראי באופן ב"ת $r_1, \dots, r_{n-1} \in \mathbb{Z}_m$ ע"י פונקציה $F(S, r) \to (s_1, \dots, s_n)$
- <u>בפתרון לעיל</u>: כל משתתף מקבל את החלק ה-s_i, כאשר $s_{i_{1 \leq i \leq n-1}} = r_i$ ו-$s_n = S - \sum_{i=1}^{n-1} r_i \ (\mathrm{mod}\ m)$ (עבור $m > 2$ חשוב מינוס ולא פלוס).

$t-out-of-n$ secret sharing:

- בדומה לקודם, מאפשרים ל-tיה לשחזר את הסוד אך לכל $(t-1)$יה לא לשחזר את הסוד או כל אינפורמציה לגביו.
- <u>הנחה נוספת</u>: U הוא שדה סופי ו-$|U| \geq n+1$.

<u>אינטרפולציית לגראנג'</u>:

- יהיו t (x_i, y_i) זוגות, ונתון כי f הוא פולינום העובר דרכם. המטרה של האינטרפולציה היא למצוא את f: $f(x) = a_{t-1}x^{t-1} + \cdots + a_i x + a_0$.
- נגדיר: $\frac{x-x_t}{x_i-x_t} \cdot \dots \cdot \frac{x-x_3}{x_i-x_3} \cdot \frac{x-x_2}{x_i-x_2} \cdot y_i$ שהם t פולינומי עזר כך ש-$f_i(x) = \begin{cases} y_i, & x = x_i \\ 0, & o/w \end{cases}$.
- נגדיר: $f(x) = \sum_{i=1}^{t} f_i(x)$ והוא הפולינום הסופי.
- אינטרפולציית פולינומים עובדת מעל כל שדה, סופי ואינסופי כיוון ש-$x_i - x_j \neq 0$ עבור $i \neq j$; ולכל אחד יש הופכי (בניגוד לחוג).

טענה: יש פולינום יחיד מדרגה $t-1$ העובר דרך הנקודות הללו. נניח בשלילה כי קיים פולינום $g \neq f$ המקיים את התנאים על כל t הנקודות, אז יש לו $t-1$ שורשים ומתקיים ב-$\mathbb{Z}_p$ ש-$f-g \equiv 0$ - פולינום האפס.

סכמת חלוקת סוד t-out-of-n של Shamir:

- יהי $S \in \mathbb{Z}_p$ הסוד.
- הדילר בוחר ערכים ראנדומליים $r_{t-1}, \dots, r_1$ ($t-1$ חלקים) מ-$\mathbb{Z}_p$.
- הדילר מגדיר פולינום עם איבר חופשי S: $f(x) = r_{t-1}x^{t-1} + \cdots + r_1 x + S$.
- $s_1, \dots, s_n$ הנשלחים למשתתפים הם ערכי $f(x)$ בנקודות המתאימות: $s_1 = f(1), \dots, s_n = f(n)$.

נכונות: כדי לשחזר: t משתתפים בונים פולינום אינטר' מהנקודות שברשותם, ומיחידות הפולינום יכולים לשחזר את $f(0) = S$. גם סודיות מתקיימת.

הערה: בהינתן t חלקים $s_{i_1}, \dots, s_{i_t}$, הסוד S ניתן לשחזור כפונקציה לינארית של החלקים: יש קבועים $b_{i_1}, \dots, b_{i_t}$ כך ש-$S = \sum_{j=1}^{t} b_{i_j} s_{i_j}$.

הטלת מטבעות בטלפון:

Hard core bits of one-way functions:

הגדרה: נאמר כי $B(x)$ הוא פרדיקט ביט-קשה ל-F חד כיוונית אם קיימת פרוצדורה P יעילה המקבלת $F(x)$ המצליחה להפוך את F תוך קריאה לפרוצדורה A המחשבת את $B(x)$ מתוך $F(x)$.

פרוטוקול אקראיות ברשת באמצעות ביטים קשים:

- אליס בוחרת $F: D \to D$ פונ' חח"ע וקשה להיפוך, ויהי $B(x)$ פרדיקט קשה ל-F.
- אליס שולחת את F, B לבוב ובוחרת $x \in D$, ומחשבת את $y = F(x)$ ו-$b = B(x)$.
- היא שולחת לבוב את y וזוהי התחייבות לערך של x.
- בוב שולח לאליס ניחוש $c \in \{0,1\}$ ל-$B(x)$. לאחר קבלת c, אליס שולחת לבוב את x וכעת הוא יכול לחשב את $b = B(x)$.
- אם $c = b$ אז בוב מנצח, אחרת אליס מנצחת.

בחירת x ושליחת $F(x)$ היא התחייבות ל-$B(x)$ (הכנסת x ל"מעטפה אטומה").

Threshold Cryptography:

- הדילר מייצר $N = p \cdot q$ עם e מפתח הצפנה (וידוא חתימה). N, e מפורסמים, d מפתח חתימה.
- תהי M הודעה ו-$h(M) = y$ הודעה "מצומצמת" לגביה רוצים לחשב $y^d (mod\ N)$ (זו החתימה).
- נהיה מעוניינים ב-$y^{s_1}, \dots, y^{s_n}$ כך שמהם ניתן להרכיב את $y^d\ mod\ N$: $y^{s_1} \cdot y^{s_2} \cdot \dots \cdot y^{s_n} = y^{\sum s_i\ mod\ \phi(N)}$.
- מה נרצה מה-S? $s_1, \dots, s_{n-1}$ יבחרו באקראי ו-$s_n = d - \sum_{i=1}^{n-1} s_i\ (mod\ \phi(N))$, והמשתתף ה-$i$ יקבל y^{s_i}.
- ב-t-out-of-n יש בעיות טכניות של אינטרפולציה במעריך (כי $(p-1)(q-1)$ הוא זוגי).

Elgammal Threshold PKC:

שינוי לסכמה המקורית של Elgammal PKC: שינוי מפתח ל-β^{2a}, אליס שולחת לבוב את $(g^{2k}, m \cdot \beta^k)$, וסכמה זו חזקה כמו המקורית. פענוח ע"י בוב: חישוב $(g^{2k})^a = (g^{2a})^k = \beta^k$.

- יהיו חלקי המפתח הפרטי $2a$ לפי שמיר: $a_1, \dots, a_n$.
- נסמן $(c_1, c_2) = (g^{2k}, m \cdot \beta^k)$, רוצים לחשב את $c_1^{a_i \bmod (p-1)}$ כדי לבצע אינטרפולציה על החזקה. כיוון שבחזקה עובדים מודולו $p-1$ ו-$\mathbb{Z}_{p-1}$ אינו שדה (כי $p-1$ זוגי – שווה $2q$), עובדים עם חזקות זוגיות של g (לכן נבחר מפתח זוגי $2a$).
- QR ב-$\mathbb{Z}_p^*$ הם תת חבורה, וכיוון ש-$p-1 = 2q$ כל תת חבורה היא בעלת או 2 או q איברים מ-QR.
- בהינתן מפתח a המחלק בונה $f[x]$ מדרגה t. כיוון ש-$c_1 = g^{2k}$ הוא שארית ריבועית, ולכן כל $c_1^{a_i}$ גם שארית ריבועית.
- עבור t כלשהם, נניח $c_1^{a_1}, \dots, c_1^{a_t}$ לפי סכמת שמיר קיימים $b_1, \dots, b_t$ כך ש-$a = \sum_{i=1}^{t} b_i a_i$ ואז:

$$m = \left(c_1^{a_1}\right)^{b_1} \cdot \dots \cdot \left(c_1^{a_t}\right)^{b_t} = c_1^{\sum_{i=1}^{t} b_i a_i} = c_1^{a} (\mathrm{mod}\, p)$$

למה threshold RSA קשה יותר:

- בהעלאה בחזקה ניתן בקלות לחבר ולכפול, ב-RSA רק לכפול.
- חזקה ב-RSA היא מודולו $(p-1)(q-1)$. נניח שהיה ניתן להפטר מחזקות של 2. לו היינו עובדים ישירות עם $\frac{(q-1)}{2} \cdot \frac{p-1}{2}$ זה מאפשר פירוק pq - ומזה רוצים להימנע. לסיכום, מסובך יותר אם כי אפשרי.

הוכחות אינטראקטיביות ו-Zero Knowledge Proofs:

רכיבי מערכת הוכחה אינטראקטיבית הנוספים על רכיבי הוכחה רגילה (אקסיומות וכו'):

- P הוא prover יודע כל ולא מוגבל חישובית.
- V הוא verifier ספקן ואינו מוגבל חישובית.
- רכיב אקראיות (הטלת מטבעות).
- מטרה: אם הטענה נכונה, V ישתכנע בהסתברות $1-\varepsilon$. אם אינה נכונה, V ישתכנע בהסתברות קטנה מ-ε.

דוגמא: Graph Isomorphism:

- P מעוניין לשכנע את V ששני גרפים G_1, G_2 הם איזומורפיים.
- V בוחר באקראי אחד משני הגרפים, מבצע תמורה על שמות הקודקודים ומחזיר את הגרף לאחר התמורה, נסמנו G ($G = \pi(G_i)$ ל-i שנבחר).
- P המוכיח צריך לחשב מה היה ה-i שנבחר, ואם $G_1 \cong G_2$, P טועה בסיכוי $\frac{1}{2}$. לאחר 200 פעם, אם P מצליח בכולם V ישתכנע, אחרת V ידחה.

Zero-Knowledge Proofs:

הרעיון: P רוצה לשכנע את V שהוא יודע להוכיח מבלי להסגיר מידע ל-V שלא יוכל לשחזר הוכחה בעצמו.

דוגמא: הוכחה שגרף הוא 3-צביע:

- אליס רוצה לשכנע את בוב שגרף הוא 3-צביע. תחילה אליס תבצע פרמוטציה על צביעת הגרף, והצפנה על הצביעה כך שלכל צומת יש מפתח משלו.
- בוב יכול לבקש הוכחה על קשת, ואליס תשלח לו את שני צבעי צמתי הקשת המוצפנים. לאחר השליחה, אליס תבצע שוב פרמוטציה על הצבעים ותצפין אותם. כיוון שנעשית פרמוטציה על הצבעים בכל שלב, בוב לא יוכל לשחזר את הצביעה.
- בכל סיבוב אם הגרף אינו 3-צביע, בוב יגלה זאת בסיכוי $\frac{1}{|E|}$. לאחר k פעמים, הסיכוי שאליס מרמה הוא $\left(1-\frac{1}{|E|}\right)^k$.

סכמות זיהוי – Identification:

מטרה: הזדהות בכדי לאפשר גישה למשאבים (למשל); אליס רוצה להזדהות בפני בוב ואיב רוצה להתחזות לאליס.

זיהוי ראשוני: Carol היא צד שלישי אמין שיכול להעביר לאליס אמצעי זיהוי ראשוניים באופן מאובטח.

סכמת פיאט-שמיר לזיהוי:

- בוב מקבל $N = pq$ מקרול אך לא את הפירוק.
- אליס בוחרת m מספרים באקראי R_i מ-$\mathbb{Z}_N$, מחשבת את הריבועים שלהם: $S_i = R_i^2$.
- אליס נותנת לבוב את כל S_i ושומרת בסוד את כל R_i.
- אליס תשכנע את בוב שהיא אכן אליס ע"י כך שתראה לו שהיא יודעת את השורשים הריבועיים של S_i מבלי לחשוף בפניו את R_i (ע"י ZK proof).

שכנוע (נסתכל על S_1):

- מתקיים כי $S_1 = R_1^2 (\mathrm{mod}\, N)$.
- אליס תבחר $X_1 \in \mathbb{Z}_N$ ומחשבת את $Y_1 = X_1^2 (mod\ N)$, ושולחת את Y_1 לבוב.
- אליס טוענת שהיא יודעת את השורש הריבועי של Y_1 ושל $Y_1 S_1$ מודולו N, ומכאן היא יודעת את השורש הריבועי של S_1.
- בוב בוחר באקראי (בהסתברות $\frac{1}{2}$) אחד מהשניים שאליס תתן לו את השורש הריבועי שלו.
- אם אליס יודעת את השורש של Y_1 ואת של $Y_1 S_1$, היא יודעת את השורש של S_1. אם היא לא יודעת את השורש של S_1, היא לא יודעת או את השורש של Y_1 או את השורש של $Y_1 S_1$ (או שניהם), ואז בהסתברות חצי בוב יעלה על מתחזה. לאחר הרבה פעמים, הסיכוי למתחזה לא להתפס קטן.
- בפרוטוקול: אליס בוחרת $Y_1, \dots, Y_m$ ובוב מטיל מטבעות $b_1, \dots, b_m$ כבחירה מבין שתי האופציות שלו. בוב מקבל רק לאחר הצלחה בכל m המקרים.

שיפור 1:

צמצום תשובת אליס לשתי תשובות במקום m תשובות, המסתמך על כך ש-$\Pi\sqrt{x_i} = \sqrt{\Pi x_i}$:

- תשובה אחת: $\Pi\sqrt{Y_i S_i}$ על כל i עבורו $b_i = 0$.
- תשובה שניה: $\Pi\sqrt{Y_i}$ על כל i עבורו $b_i = 1$.

טענה: מול אליס, איב לא יכולה ללמוד שום דבר חדש.

מרכיבים חיוניים: אינטראקציה וראנדומיזציה.

שיפור 2:

בוב מעביר לאליס במקום הגרלת b_i שונים את $b_1 b_2 \dots b_m = H(Y_1, \dots, Y_m)$ - פוני hash על Y_i שקיבל מאליס. אבל, הפוני H ידועה לכולם והיא פסאודו-אקראית (פוני hash מאובטחת).

מערכות עם מפתח פומבי מסוג knapsack:

בעיית $subset - sum \in NPC$: נתונים n טבעיים שונים $a_1, \dots, a_n > 0$ ונתון $S \in \mathbb{N}$ (כולם בייצוג בינארי). השאלה (הכרעה): האם יש תת קבוצה של ה-a_i שסכומה הוא S. בעיה זו שקולה לבעיה: האם קיים וקטור $x_1 \dots x_n \in \{0,1\}^n$ כך ש-$\sum_{i=1}^{n} x_i a_i = S$.

מערכת מסוג knapsack:

- מפתח פומבי: $a_1, \dots, a_n$.
- הצפנה: $ciphertext = S = \sum_{i=1}^{n} x_i a_i$ כך ש-$x_1 \dots x_n \in \{0,1\}^n$

כדי לאפשר פענוח מוטב שיהיה מבנה נסתר (trapdoor info) שיתן יתרון למקבל הלגיטימי על פני המאזין. נגדיר סדרה super increasing המקיימת:

- $a_2 > a_1$
- $a_3 > a_2 + a_1$
- $a_i > \sum_{j=1}^{i-1} a_j$

למשל (לא לקריפטו): $a_i = 2^i$. טענה: אם $\{a_i\}$ סדרה סופר עולה, הפענוח קל (בדיקה בינארית על כל איברי הסדרה האם הם נכנסים לסכום).

הסתרת המבנה הסופר עולה לפי MH: נבחר M הגדול מסכום ה-b_i שהם איברי הסדרה הסופר עולה הראשונית שמתחילים איתה. נבחר $W < M$ וזר ל-M, ונגדיר $a_i' = b_i \cdot W (mod\ M)$. נבחר באקראי תמורה $\pi \in S_n$ (תמורה על n איברים) ונגדיר $a_i = a'_{\pi(i)}$ (תמורה של ה-a_i') ומפרסמים את $\{a_i\}$.

- מפתח פומבי: $a_1, \dots, a_n$
- מפתח פרטי: M (מודולוס), W (כופל) ו-π (התמורה).
- פענוח: נתון $S = \sum_{i=1}^{n} x_i a_i$, x_i לא ידועים ורוצים למצוא אותם.

כיוון ש-W זר ל-M, יש לו הופכי $W^{-1} (mod\ M)$. נחשב: $C = S \cdot W^{-1} = \sum_{i=1}^{n} x_i (a_i W^{-1})\ (mod\ M) = \sum_{i=1}^{n} x_i b_{\pi(i)}$

בעזרת תכונת הסדרה הסופר עולה מוצאים איזה $x_i = 1$ ולבסוף מפעילים את התמורה ההופכית למצוא מיקומים מקוריים. מומלץ לבחור את $b_i \approx 2^{i+n}$.

שריגים מעל $\mathbb{Z}^n$:

נתון בסיס $\overrightarrow{v_1}, \dots, \overrightarrow{v_n} \in \mathbb{Z}^n$ (n וקטורים שכל הקורדינטות שלהם שלמות). השריג הנפרש ע"י $\{\overrightarrow{v_i}\}$ הוא $L(\overrightarrow{v_1}, \dots, \overrightarrow{v_n}) = \{\sum_{i=1}^{n} a_i \overrightarrow{v_i} \mid a_i \in \mathbb{Z}\}$.

שאלה בעלת עניין (וקשה): מהו הוקטור הקצר ביותר השונה מ-0 הנפרש ע"י הבסיס?

אלגוריתם של LLL: אם הוקטור הקצר ביותר בשריג n מימדי קצר בהרבה מהשני – יש אלג' יעיל למצוא אותו.

ולענייננו: שבירת knapsack "גנרית": בהינתן מערכת knapsack נבנה עבורה שריג $n+1$ מימדי:

$$\begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \dots & & \ddots & \dots & \dots \\ 0 & 0 & \dots & 1 & -a_n \\ 0 & 0 & \dots & 0 & S \end{bmatrix}$$

וקטור הקשור לפענוח: $\sum_{i=1}^{n} x_i v_i + v_{n+1} = (x_1, x_2, \dots, x_n, 0)$ כאשר קור' ה-x_i הם 0 או 1. אם a_i גדולים יחסית, הוקטור הנ"ל יהיה קצר מאוד ואלגוריתם LLL ימצא אותו ואת ה-x_i ולכן יפענח. היוריסטיקה (לא מוכח ב-100%) שעובדת!