به نام آن که هیچ رمزی برایش پوشیده نیست

تمرین سری ۱ رمزنگاری

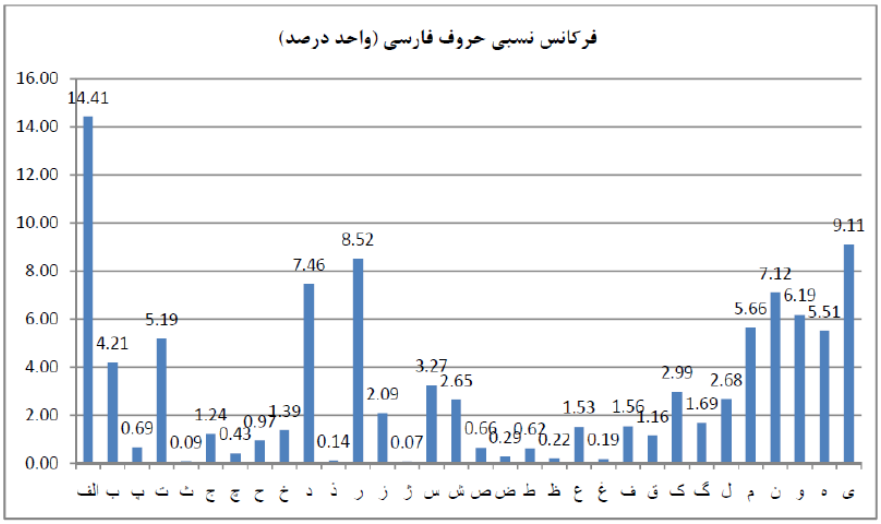
دانشکده ریاضی، آمار و علوم کامپیوتر

ترم اول سال تحصیلی ۹۵-۹۴

تاریخ تحویل: دوشنبه ۹۴/۰۷/۲۷

۱. به کمک نمودار فرکانس نسبی یا توزیع احتمال حروف فارسی که در شکل ۱ نشان داده شده، متن زیر را رمزگشایی کرده و کلید را بدست آورید.

شم فکخژد غذف کژ ژگ ثکژد سغذژرزف فک مغصزرژ ژگ ثکژد زژکز سیحسزف جر فک ثکژد فکخژد سگکچصکزد سزخژکزرژ زبدز جتک ذ دتژث ذ مکجنز ذ چخکژرز ژمص شم سرذمزحر ثکژد غذژمصررژز غذف کژ ژگ غفژ سغذژرزف ذ سژ فذمصز ثکژد سر غفژ کذز ژذکزف



شکل ۱: فرکانس نسبی حروف فارسی

۲. متن رمز شده زیر با استفاده از سیستم رمزنگاری ویجنر [1] رمز شده، کلید و متن اصلی را بدست آورید.

CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWI
AKLXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJEL
XVRVPRTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJTAM
RVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBIPEE
WEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHPWQ
AIIWXNRMGWOIIFKEE

الگوریتم رمز ویجنر به صورت زیر کار می‌کند:
فرض کنید که m یک عدد طبیعی باشد. الفبای متن اصلی و رمز شده و الفبای کلید را $\mathbb{Z}_{۲۶}$ که معادل با همان کاراکترهای زبان لاتین است در نظر بگیرید. این سیستم رمزنگاری یک سیستم متقارن است که کلید آن یک m تایی از اعداد در $\mathbb{Z}_{۲۶}$ (کاراکترهای لاتین) است. یعنی $K = (k_۱, k_۲, ..., k_m)$ برای رمزنگاری متن x بلوک‌های m تایی از آن مثل $(x_۱, x_۲, ..., x_m)$ را گرفته و به صورت زیر رمز می‌کنیم تا کل متن به این ترتیب رمز شود.

$$e_k(x_۱, x_۲, ..., x_m) = (x_۱ + k_۱, x_۲ + k_۲, ..., x_m + k_m)$$

[1] Vigenere

برای رمزگشایی متن رمز شده y به صورت زیر عمل می کنیم.

$$d_k(y_1, y_2, ..., y_m) = (y_1 - k_1, y_2 - k_2, ..., y_m - k_m)$$

همه‌ی جمع و تفریق‌های فوق در پیمانه ۲۶ است. بنابراین برای این که بتوانید متن رمز شده فوق را رمزگشایی کنید ابتدا باید طول کلید یعنی m را بیابید و سپس m کاراکتر کلید را بدست آورده و متن را رمزگشایی کنید. راهنمایی: بهتر است از ویژگی‌های آماری زبان لاتین به خصوص تک حرفی‌ها استفاده کنید. برای مشاهده ویژگی‌های آماری زبان لاتین به لینک زیر مراجعه کنید: فرکانس نسبی حروف لاتین

۳. فرض کنید که p یک عدد اول باشد.

(آ) نشان دهید تعداد ماتریس‌های معکوس پذیر ۲ × ۲ روی $\mathbb{Z}_p$ برابر است با:

$$(p^2 - 1)(p^2 - p)$$

راهنمایی: چون p اول است پس $\mathbb{Z}_p$ میدان است، از این رو می‌توانید از این حقیقت که یک ماتریس روی یک میدان معکوس‌پذیر است اگر و تنها اگر تمام ستون‌های آن مستقل خطی باشند استفاده کنید.

(ب) در زیر یک سیستم رمزنگاری مشابه با سیستم هیل [۲] شرح داده شده، اندازه فضای کلید آن را بدست آورید.

$$m \geq 2; \mathcal{M} = \mathcal{C} = \mathbb{Z}_p^m \qquad \mathcal{K} = \{K \in \mathbb{Z}_p^{m \times m} : K\text{معکوس پذیر باشد}\}$$

یعنی فضای متن اصلی و رمزشده بردارهای ستونی با m درایه از اعضای $\mathbb{Z}_p$ هستند و فضای کلید شامل همه‌ی ماتریس‌های $m \times m$ معکوس پذیر با درایه‌هایی از $\mathbb{Z}_p$ است. رمزنگاری و رمزگشایی نیز به صورت زیر انجام می‌شود.

$$e_K(x) = xK$$

$$d_K(y) = yK^{-1}$$

همه‌ی اعمال فوق در پیمانه‌ی p انجام می‌شود.

۴. سیستم رمز متقارنی با فضای پیام $\mathcal{M}$ و فضای متن رمز شده $\mathcal{C}$ را در نظر بگیرید به طوری که برای هر زوج متن اصلی مثل $m., m_1 \in \mathcal{M}$ و هر متن رمز شده $c \in \mathcal{C}$ داشته باشیم:

$$pr\{C = c|M = m.\} = pr\{C = c|M = m_1\}$$

در رابطه فوق M و C به ترتیب متغیرهای تصادفی متن اصلی و متن رمز شده هستند. از این جهت می‌گوییم متغیر تصادفی چون متن اصلی می‌تواند به طور تصادفی از فضای متن اصلی انتخاب شود و در این صورت متن رمز شده که به وسیله یک الگوریتم قطعی از روی متن اصلی بدست می‌اید، نیز یک متغیر تصادفی است. به بیان شهودی، اگر $m.$ و m_1 را رمز کنیم، هر یک با احتمال برابری ممکن است به c رمز شوند و ما نمی‌توانیم بین این دو تمایزی قائل شویم. چنین سیستمی را امن کامل گوییم. [۳] نشان دهید در چنین سیستم رمزنگاری حتما اندازه فضای کلید یعنی $|\mathcal{K}|$ باید بزرگتر یا مساوی اندازه فضای متن اصلی یعنی $|\mathcal{M}|$ باشد.

۵. سناریوی زیر را در نظر بگیرید:
بهرام و آذر می‌خواهند آزمایش پرتاب سکه را انجام دهند. اگر شیر آمد آذر و در غیر این صورت بهرام برنده بازی است. ولی مشکل در این جا است که این دو به یکدیگر دسترسی ندارند و فقط می توانند از طریق تلفن با هم ارتباط داشته باشند!
پروتوکولی را طراحی کنید که آن‌ها بتوانند این آزمایش را انجام دهند طوری که هیچ کدام از آن دو نتوانند تقلب کنند. (راهنمایی: به فصل ۱ کتاب Mihir Bellare مراجعه کنید.)

۶. معمولا برای ذخیره سازی یا انتقال داده‌ها از فشرده سازی استفاده می‌شود. برای مثال در فشرده کردن داده‌های تصویری از استاندارد jpeg استفاده می‌شود. فرض کنید بخواهیم یک داده را ضمن فشرده‌سازی، رمزنگاری کنید یعنی به صورت ترکیبی از رمزنگاری و فشرده سازی استفاده کنیم. آیا بهتر است ابتدا رمزنگاری کنیم و سپس حاصل را فشرده کنیم یا بالعکس؟(دلیل بیاورید)

۷. نشان دهید برای همه‌ی کلیدهای $K \in \{0,1\}^{56}$ و برای همه‌ی پیام‌های $x \in \{0,1\}^{64}$ داریم:

$$DES_K(x) = \overline{DES}_{\overline{K}}(\overline{x})$$

این ویژگی DES را key-complementation می‌گویند.

۸. توضیح دهید که ویژگی key-complementation چگونه حمله جست‌وجوی کامل فضای کلید [۴] را از مرتبه ۲ کاهش می‌دهد، یعنی کافی است نیمی از کل فضای کلید را جست‌و جو کنید!.

---

[۲] Hill

[۳] C. E. Shannon: A mathematical theory of communication. Bell System Technical Journal, vol. 27,pp. 379–423 and 623–656, July and October, 1948

[۴] brute force attack

۹. (آ) یک سبک استفاده از الگوریتم‌های رمز قالبی، سبک شمارنده (counter mode) است در این سبک، الگوریتم رمز قالبی به یک الگوریتم رمز دنباله‌ای تبدیل می‌شود. و بیت‌های کلید به صورت زیر تولید می‌شوند:

$$z = (E_K(۰), E_K(۱), E_K(۲), ...)$$

فرض کنید الگورریتم رمز قالبی $E_K(.)$ همان الگوریتم DES باشد(با کلید ۵۶ بیتی و پیام ۶۴ بیتی). روشی مشخص کنید که با احتمال بالایی بتواند تشخیص دهد که یک دنباله داده شده، توسط الگوریتم DES و با سبک فوق تولید شده یا این که یک دنباله‌ی تصادفی است.(دنباله داده شده شامل N بلوک ۶۴ بیتی است و N دلخواه است، به عبارت دیگر طول دنباله‌ی باینری داده شده مضربی از طول قالب یا همان ۶۴ است.)

راهنمایی: با توجه به این که $E_k(.)$ یک به یک است $(x \neq y \Rightarrow E_K(x) \neq E_k(y))$ لذا اگر دنباله کلید توسط DES تولید شده باشد، قالب‌های ۶۴ بیتی دنباله هرگز نباید تکرار شوند(تا قبل از اتمام یک دوره تناوب) در حالی که در یک دنباله کاملا تصادفی امکان تکرار این قالب‌ها وجود دارد. از این ویژگی استفاده کرده و یک روش برای تمایز دنباله تولید شده توسط DES در سبک شمارنده از یک دنباله‌ی کاملا تصادفی ارائه دهید.

(ب) چه طولی از دنباله داده شده لازم است تاتمایزگر پیشنهادی شما با احتمال بیش از $\frac{۱}{۲}$ بتواند تمایز دهد؟(در صورت لزوم از تقریب $e^{-x} = ۱ - x$ استفاده کنید.

۱۰. در میدان $GF(۲^۸)$ با چند جمله‌ای تحویل ناپذیر $m(x) = x^۸ + x^۴ + x^۳ + x + ۱$ حاصل ضرب زیر را حساب کنید.

$$\{e۱\}\{۰۵\}$$

یادآوری: تناظر زیر بین بایت‌ها و اعضای $GF(۲^۸)$ وجود دارد:

$$a = a_۷a_۶a_۵a_۴a_۳a_۲a_۱a_۰. \longleftrightarrow m(x) = a_۷x^۷ + a_۶x^۶ + a_۵x^۵ + a_۴x^۴ + a_۳x^۳ + a_۲x^۲ + a_۱x + a_۰.$$

۱۱. یکی از مراحلی که در دورهای میانی الگوریتم AES انجام می‌شود، مرحله $mix - cols$ است. این مرحله را می‌توان به دو صورت زیر بیان کرد:

(آ) فرض کنید بخواهید بردار $\begin{bmatrix} a_۰ \\ a_۱ \\ a_۲ \\ a_۳ \end{bmatrix}$ را از مرحله mix-cols عبور دهیم. می توان این بردار را به عنوان چند جمله‌ای $a(x) = a_۳x^۳ + a_۲x^۲ + a_۱x + a_۰$ در نظر گرفت و آن را در چند جمله‌ای زیر(که یک چند جمله‌ای قراردادی برای این الگوریتم است) در پیمانه‌ی $x^۸ + ۱$ ضرب کرد تا حاصل mix شده بدست آید.

$$c(x) = \{۰۳\}x^۳ + \{۰۱\}x^۲ + \{۰۱\}x + \{۰۲\}$$

(ب) روش دیگر این است که از ضرب ماتریسی زیر برای mix کردن این ستون استفاده کنیم:

$$\begin{bmatrix} b_۰ \\ b_۱ \\ b_۲ \\ b_۳ \end{bmatrix} = \begin{bmatrix} ۰۲ & ۰۳ & ۰۱ & ۰۱ \\ ۰۱ & ۰۲ & ۰۳ & ۰۱ \\ ۰۱ & ۰۲ & ۰۲ & ۰۳ \\ ۰۳ & ۰۱ & ۰۱ & ۰۲ \end{bmatrix} \begin{bmatrix} a_۰ \\ a_۱ \\ a_۲ \\ a_۳ \end{bmatrix}$$

دقت کنید که ضرب ماتریسی فوق باید در میدان $GF(۲^۸)$ انجام شود. **نشان دهید این دو روش معادل‌اند.**