به نام خدا

دانشکده‌ی علوم ریاضی

مقدمه‌ای بر رمزنگاری ۲۷ اردیبهشت ۱۳۹۲

جلسه‌ی ۱۳: امنیت CPA و CCA، جایگشت شبه‌تصادفی

مدرس: دکتر شهرام خزائی *نگارنده: نیلوفر صفی صمغ آبادی*

۱ تابع شبه‌تصادفی و رمز دنباله‌ای

در جلسه قبل RF_n را مجموعه‌ی همه توابعی در نظر گرفتیم که n بیت را به n بیت می‌نگارند و همان‌طور که مطرح کردیم تعداد کل چنین توابعی $\mathbf{2}^{n\mathbf{2}^n}$ است. تابعی که به طور کاملاً تصادفی از این مجموعه انتخاب شده باشد، یک تابع تصادفی نامیده می‌شود. در عمل به جای استفاده از توابع تصادفی، از توابع شبه‌تصادفی استفاده می‌کنیم زیرا برای نمایش و کد کردن یک تابع تصادفی به $n\mathbf{2}^n = \log(\mathbf{2}^{n\mathbf{2}^n})$ بیت نیاز است که حتی برای مقادیر نسبتاً کوچک n، نمی‌توانیم از این توابع بطور عملی استفاده کنیم. یک تابع شبه‌تصادفی نامیده می‌شود اگر هیچ مهاجم کارایی نتواند با دسترسی اوراکلی به تابع، آن را از یک تابع کاملاً تصادفی با مزیت قابل توجهی تمیز دهد. برای دقیق کردن این مفهوم، می‌توان از رویکرد دقیق[۱] و مجانبی[۲] استفاده کرد. ما از رویکرد دوم استفاده کردیم که برای این منظور لازم است تعریف برای خانواده‌ای از توابع مطرح شود.

تعریف ۱ (تابع شبه‌تصادفی) *خانواده‌ی توابع $\{f_k : \{\mathbf{0},\mathbf{1}\}^{|k|} \to \{\mathbf{0},\mathbf{1}\}^{|k|}\}_{k\in\{\mathbf{0},\mathbf{1}\}^*}$ را شبه‌تصادفی[۳] گوییم اگر:*

- *یک الگوریتم چندجمله‌ای وجود داشته باشد که بتواند $f_k(x)$ را از روی k و x محاسبه کند،*
- *تابع ناچیز $\varepsilon(n)$ وجود داشته باشد که:*

$$|\Pr\{f \leftarrow \mathsf{RF}_n : \mathcal{A}^{f(\cdot)}(\mathbf{1}^n) = \mathbf{1}\} - \Pr\{k \leftarrow \{\mathbf{0},\mathbf{1}\}^n : \mathcal{A}^{f_k(\cdot)}(\mathbf{1}^n) = \mathbf{1}\}| \leq \varepsilon(n)$$

یک روش طراحی توابع شبه‌تصادفی در عمل، استفاده از رمزهای دنباله‌ای مدرن است که دارای IV هستند. با استفاده از یک خانواده از توابع شبه‌تصادفی، می‌توان سیستم رمز دنباله‌ای به صورت زیر طراحی کرد.

تعریف ۲ (رمز دنباله‌ای) *فرض کنید $\{f_k : \{\mathbf{0},\mathbf{1}\}^{|k|} \to \{\mathbf{0},\mathbf{1}\}^{|k|}\}_{k\in\{\mathbf{0},\mathbf{1}\}^*}$ یک خانواده از توابه شبه‌تصادفی باشد. سیستم رمز $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ روی فضای پیام $\mathcal{M} = \{\mathbf{0},\mathbf{1}\}^n$ که به شکل زیر تعریف می‌شود یک رمز دنباله‌ای است:*

---

[۱] concrete
[۲] asymptotic
[۳] pseudo-random

- **الگوریتم تولید کلید:** *یک رشته تصادفی n-بیتی را به عنوان کلید تولید می‌کند:*

$$\mathsf{Gen}(1^n) : k \leftarrow \{0,1\}^n$$

- **الگوریتم رمزنگاری:** *برای رمز کردن پیام $m \in \{0,1\}^n$ تحت کلید $k \in \{0,1\}^n$، ابتدا رشته‌ی تصادفی $r \leftarrow \{0,1\}^n$ را تولید کرده و سپس $\langle r, f_k(r) \oplus m \rangle$ را به عنوان متن رمزی به خروجی می‌دهد:*

$$c \leftarrow \mathsf{Enc}_k(m) : r \leftarrow \{0,1\}^n, \langle r, c \rangle \leftarrow \langle r, f_k(r) \oplus m \rangle$$

- **الگوریتم رمزگشایی:** *برای یافتن پیام m متناظر با متن‌رمزی $c = \langle r, u \rangle$ که $r, u \in \{0,1\}^n$ به صورت زیر عمل می‌کند:*

$$m \leftarrow \mathsf{Dec}_k(\langle r, u \rangle) : m = u \oplus f_k(r)$$

۲ تعاریف مختلف امنیت

تاکنون با تعریف امنیت یک‌پیامی و چندپیامی[۴] آشنا شده‌ایم که صرفاً مهاجم شنودگر را مدل می‌کند. در این بخش پس از یادآوری امنیت چندپیامی، امنیت‌های قوی‌تر متن اصلی انتخابی (CPA[۵]) و متن رمزی انتخابی (CCA[۶]) را معرفی می‌کنیم. امنیت CPA مهاجمی را مدل می‌کند که به دستگاه رمزنگاری دسترسی دارد. امنیت CCA مهاجمی را مدل می‌کند که علاوه‌بر دسترسی به دستگاه رمزنگاری، دستگاه رمزگشایی را نیز در اختیار دارد.

۱.۲ امنیت چندپیامی

تعریف ۳ (آزمایش امنیت چندپیامی $\mathsf{PrivK}^{\mathrm{mult}}_{\mathcal{A},\Pi}(n)$) *مراحل آزمایش امنیت تک‌پیامی به صورت زیر است:*

۱. $k \leftarrow \mathsf{Gen}(1^n)$

۲. $(m_0^1, \ldots, m_0^{p(n)}), (m_1^1, \ldots, m_1^{p(n)}) \leftarrow \mathcal{A}(1^n)$ که $|m_0^i| = |m_1^i|$

۳. $b \leftarrow \{0,1\}$

۴. $c_i \leftarrow \mathsf{Enc}_k(m_b^i)$ *برای* $i = 1, \ldots, p(n)$.

۵. $\widehat{b} \leftarrow \mathcal{A}(c_1, \ldots, c_{p(n)})$

خروجی این آزمایش، با متغیر تصادفی $\mathsf{PrivK}^{\mathrm{mult}}_{\mathcal{A},\Pi}(n)$ نشان داده می‌شود و برابر ۱ تعریف می‌شود اگر $\widehat{b} = b$ باشد؛ که احتمال رخداد آن، هم به سیستم رمز و هم به استراتژی مهاجم[۷] بستگی دارد.

[۴] multi-message security
[۵] chosen plaintext attack security
[۶] chosen ciphertext attack security
[۷] adversary

تعریف ۴ (امنیت چندپیامی) *سیستم رمز* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *دارای امنیت چندپیامی است، اگر برای هر مهاجم چندجمله‌ای- زمان تصادفی غیریکنواخت (*$nuPPT$[۸]*) مانند* $\mathcal{A}$*، تابع ناچیز* $\varepsilon(n)$ *موجود باشد که:*

$$\Pr\{\mathsf{PrivK}^{\mathrm{mult}}_{\mathcal{A},\Pi}(n) = ۱\} \leq \frac{۱}{۲} + \varepsilon(n)$$

قضیه ۱ *سیستم رمز معرفی شده در تعریف ۲ ، دارای امنیت چندپیامی است.*

در ادامه به اثبات شهودی این موضوع می‌پردازیم ولی از اثبات دقیق آن صرف‌نظر می‌کنیم.
آزمایش امنیت چندپیامی را درنظر بگیرید. چنانچه مهاجم در این آزمایش دسته‌ی اول پیام‌ها را برگزیند، چیزی که دریافت می‌کند

$$\langle r_۱, m^{۱}_{۰} \oplus f_k(r_۱)\rangle, \ldots, \langle r_{p(n)}, m^{p(n)}_{۰} \oplus f_k(r_{p(n)})\rangle$$

و در صورتی که دسته‌ی دوم را برگزیند،

$$\langle r_۱, m^{۱}_{۱} \oplus f_k(r_۱)\rangle, \ldots, \langle r_{p(n)}, m^{p(n)}_{۱} \oplus f_k(r_{p(n)})\rangle$$

است. عامل اصلی که موجب می‌شود مهاجم نتواند دو دسته متن رمزشده‌ی مذکور را از هم تشخیص دهد این است که چون تابع f شبه‌تصادفی است، مهاجم نمی‌تواند مقادیر $f_k(r_i)$ ها را از مقادیر تصادفی تشخیص دهد. البته اگر مهاجم شناس داشته باشد و دو تا از مقادیر r_i ها یکسان باشند، قادر به تشخیص متن‌های رمزشده از یکدیگر خواهد بود. اما احتمال چنین رویدادی برابر $\frac{\binom{p(n)}{۲}}{۲^n}$ است که برای هر تابع چندجمله‌ای $p(\cdot)$ یک احتمال ناچیز است. طبق قضیه روز تولد، این احتمال برای $p(n) \approx ۲^{n/۲}$ احتمال قابل توجهی است که برای مهاجم‌های چندجمله‌ای قابل استفاده نیست.

۲.۲ امنیت متن اصلی انتخابی

تعریف ۵ (آزمایش حمله‌ی متن اصلی انتخابی $\mathsf{PrivK}^{\mathrm{cpa}}_{\mathcal{A},\Pi}(n)$**)** *در آزمایش متن اصلی انتخابی، مهاجم به الگوریتم رمزنگاری دسترسی اوراکلی*[۹] *دارد پ مراحل آن به صورت زیر است:*

۱. $k \leftarrow \mathsf{Gen}(۱^n)$

۲. $m_۰, m_۱ \leftarrow \mathcal{A}^{\mathsf{Enc}_k(\cdot)}(۱^n)$ که $|m_۰| = |m_۱|$

۳. $b \leftarrow \{۰, ۱\}$

۴. $c \leftarrow \mathsf{Enc}_k(m_b)$

۵. $\widehat{b} \leftarrow \mathcal{A}^{\mathsf{Enc}_k(\cdot)}(c)$

خروجی این آزمایش، با متغیر تصادفی $\mathsf{PrivK}^{\mathrm{cpa}}_{\mathcal{A},\Pi}(n)$ *تعریف می‌شود و مقدار آن برابر ۱ است اگر* $\widehat{b} = b$*.*

[۸] non-uniform probabilistic polynomial time
[۹] oracle access

نکته ۱ *چون الگوریتم‌های رمزنگاری شرکت‌کننده در آزمایش بالا می‌توانند احتمالاتی[۱۰] باشند، اگر مهاجم به تعداد چندجمله‌ای بار بخواهد پیامی را رمز کند، هر بار می‌تواند متن رمزشده‌ی متفاوتی دریافت کند.*

تعریف ۶ (امنیت CPA) *سیستم رمز* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *دارای امنیت متن اصلی انتخابی (یا* CPA*-امن[۱۱]) است، اگر برای هر مهاجم چندجمله‌ای- زمان تصادفی غیریکنواخت مانند* $\mathcal{A}$*، تابع ناچیز* $\varepsilon(n)$ *وجود داشته باشد که:*

$$\Pr\{\mathsf{PrivK}^{\mathrm{cpa}}_{\mathcal{A},\Pi}(n) = ۱\} \leq \frac{۱}{۲} + \varepsilon(n)$$

نکته ۲ *یک روش عملی ساخت سیستم‌های رمزی* CPA*-امن، استفاده از رمزهای دنباله‌ای[۱۲] مدرن است که علاوه بر کلید، از یک مقدار اولیه (*IV*[۱۳]) نیز استفاده می‌کند.*

قضیه ۲ *سیستم رمز معرفی شده در تعریف ۲ ،* CPA*-امن است.*

۳.۲ امنیت متن رمزی انتخابی

تعریف ۷ (آزمایش حمله‌ی متن رمزی انتخابی $\mathsf{PrivK}^{\mathrm{cca}}_{\mathcal{A},\Pi}(n)$**)** *در آزمایش متن رمزی انتخابی، مهاجم علاوه بر الگوریتم رمزنگاری، به الگوریتم رمزگشایی نیز دسترسی اوراکلی دارد و مراحل آن به صورت زیر است:*

۱. $k \leftarrow \mathsf{Gen}(۱^n)$

۲. $m_۰, m_۱ \leftarrow \mathcal{A}^{\mathsf{Enc}_k(\cdot),\mathsf{Dec}_k(\cdot)}(۱^n)$ که $|m_۰| = |m_۱|$

۳. $b \leftarrow \{۰, ۱\}$

۴. $c \leftarrow \mathsf{Enc}_k(m_b)$

۵. $\widehat{b} \leftarrow \mathcal{A}^{\mathsf{Enc}_k(\cdot),\mathsf{Dec}_k(\cdot)}(c)$ که $\mathcal{A}$ مجاز به درخواست رمزگشایی c نیست.

خروجی آزمایش، با متغیر تصادفی $\mathsf{PrivK}^{\mathrm{cca}}_{\mathcal{A},\Pi}(n)$ *تعریف می‌شود و مقدار آن برابر ۱ است، در صورتیکه* $\widehat{b} = b$*.*

تعریف ۸ (امنیت CCA) *سیستم رمز* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *دارای امنیت متن رمزی انتخابی (یا* CPA*-امن[۱۴]) است، اگر برای هر مهاجم چندجمله‌ای- زمان تصادفی غیریکنواخت مانند* $\mathcal{A}$*، تابع ناچیز* $\varepsilon(n)$ *وجود داشته باشد که:*

$$\Pr\{\mathsf{PrivK}^{\mathrm{cpa}}_{\mathcal{A},\Pi}(n) = ۱\} \leq \frac{۱}{۲} + \varepsilon(n)$$

[۱۰] probabilistic
[۱۱] chosen plaintext attack secure
[۱۲] stream cipher
[۱۳] initial value
[۱۴] chosen ciphertext attack secure

مثال ۹ *سیستم رمز معرفی شده در تعریف ۲، CCA-امن نیست.*
برای اثبات این موضوع، کافیست نشان دهیم مهاجم تصادفی چندجمله‌ای $\mathcal{A}$ وجود دارد که می‌تواند در آزمایش حمله‌ی متن رمزی انتخابی با احتمال $\frac{1}{2} + \mu(n)$ موفق شود که $\mu(n)$ یک تابع غیرناچیز (قابل توجه) است. مهاجم $\mathcal{A}$ را به این صورت می‌سازیم که در مرحله دوم آزمایش (بدون استفاده از اوراکل‌های رمزنگاری و رمزگشایی) دو دسته پیام تمام صفر $m_0 = 0^n$ و تمام یک $m_1 = 1^n$ را تولید کرده و در اختیار چالشگر[15] قرار می‌دهد. مهاجم منتظر می‌ماند تا چالشگر بیت تصادفی b را تولید و پیام m_b را توسط الگوریتم رمزنگاری رمز نموده و پیام رمزی $c = \langle r, u\rangle$ را در اختیار او قرار دهد. سپس در مرحله ۵ آزمایش، مهاجم با دریافت متن رمزی $c = \langle r, u\rangle$ بدین گونه عمل می‌کند: با توجه به اینکه مهاجم به توابع رمزنگار و رمزگشا دسترسی اوراکلی دارد، مهاجم پیام متناظر با متن رمزی $c' = \langle r, u \oplus 10^{n-1}\rangle$ را از تابع رمزگشا درخواست می‌کند و پس از دریافت پیامِ متناظر آن، $m' = \mathsf{Dec}_k(c')$، بیت $\widehat{b}$ را بصورت زیر تولید می‌کند:

$$\widehat{b} = \begin{cases} 0 & m' = 10^{n-1} \\ 1 & m' = 01^{n-1} \end{cases}$$

دقت کنید که چون $c' \neq c$ مهاجم مجاز به درخواست رمزگشایی c' است. با توجه به اینکه $u = m \oplus f_k(r)$ و $u \oplus 10^{n-1} = m' \oplus f_k(r)$ داریم $m' = 10^{n-1} \oplus m_b$. بنابراین همواره $b = \widehat{b}$ و لذا احتمال موفقیت مهاجم برابر است با:

$$\Pr\{\mathsf{PrivK}^{\mathrm{cca}}_{\mathcal{A},\Pi}(n) = 1\} = 1$$

که ناچیز نیست. لذا سیستم رمز مورد نظر CCA-امن نیست.

۳ جایگشت شبه‌تصادفی و رمزقالبی

تعریف ۱۰ (جایگشت) *تابع $E : \{0,1\}^n \to \{0,1\}^n$ را یک جایگشت[16] گوییم، اگر تابع معکوس*

$$E^{-1} : \{0,1\}^n \to \{0,1\}^n$$

موجود باشد به طوری که:

$$\forall x \in \{0,1\}^n : E^{-1}(E(x)) = x$$

به تعبیر دیگر تابع E را جایگشت می‌گوییم، اگر و تنها اگر یک به یک باشد. مشابه توابع شبه‌تصادفی می‌توان مفهوم شبه‌تصادفی بودن را برای جایگشت‌ها نیز مطرح کرد. یک جایگشت شبه‌تصادفی نامیده می‌شود اگر هیچ مهاجم کارایی نتواند با دسترسی اوراکلی به جایگشت و معکوس‌اش، آن را از یک جایگشت کاملاً تصادفی با مزیت قابل توجهی تمیز دهد. تعریف را می‌توان به صورت دقیق و یا مجانبی ارائه کرد که ما از رویکر دوم استفاده می‌کنیم.

تعریف ۱۱ (جایگشت شبه‌تصادفی) *خانواده‌ی جایگشت‌های*

[15] challenger
[16] permutation

$$\{E_k : \{۰,۱\}^{|k|} \to \{۰,۱\}^{|k|}\}_{k\in\{۰,۱\}^*}$$

را شبه‌تصادفی گوییم اگر:

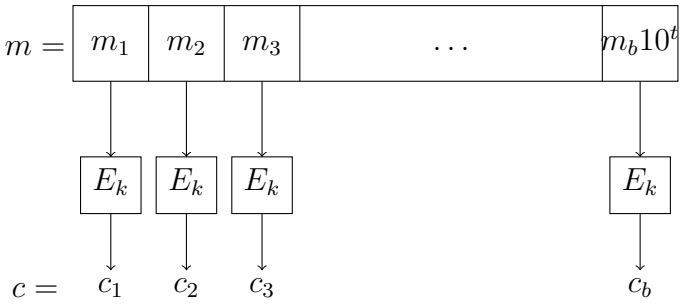
- *یک الگوریتم چندجمله‌ای وجود داشته باشد که بتواند $E_k(m)$ را از روی k و m محاسبه کند.*
- *برای هر مهاجم $\mathcal{A}$ تابع ناچیز $\varepsilon(n)$ وجود داشته باشد که:*

$$|\Pr\{\mathcal{A}^{E(\cdot),E^{-۱}(\cdot)}(۱^n)=۱\} - \Pr\{k\leftarrow\{۰,۱\}^n : \mathcal{A}^{E_k(\cdot),E_k^{-۱}(\cdot)}(۱^n)=۱\}| \le \varepsilon(n)$$

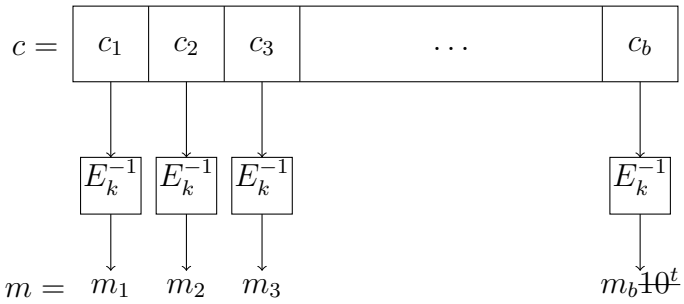
که RP_n مجموعه همه جایگشت‌های روی رشته‌های n-بیتی است و اندازه آن $!۲^n$ است.

گزاره ۱ *اگر E یک جایگشت شبه‌تصادفی باشد، یک تابع شبه‌تصادفی نیز می‌باشد.*

جایگشت‌های شبه‌تصادفی کاربردهای بسیار زیادی در رمزنگاری دارد و تحت عنوان رمزقالبی[۱۷] شناخته می‌شوند. از رمزهای قالبی معروف می‌توان به DES و AES اشاره کرد. این جایشگت‌ها برای طول ورودی (قالب) و طول کلید خاص طراحی شده‌اند. رمز قالبی DES دارای طول قالب ۶۴ و طول کلید ۵۶ بیت است. رمز قالبی AES دارای سه نسخه AES-128، AES-192 و AES-256 است که عدد بکار رفته بیانگر طول کلید است؛ طول قالب هر سه نسخه ۱۲۸ بیت است. طول قالب کوچک DES آن را مناسب برای استفاده به عنوان یک تابع شبه‌تصادفی نمی‌کند زیرا با استفاده از قضیه روز تولد با $۲^{۳۲}$ پرسمان از یک تابع تصادفی قابل تمیز است. در زمان طراحی DES طول کلید ۵۶ مناسب بود اما امروزه این طول سطح مناسبی از امنیت برآورده نمی‌کند، به طوری که با هزینه چندهزار دلار می‌توان فضای کلید را در چند روز جستجو کرد.
رمزهای قالبی برای اعمال روی قالب‌های داده طراحی شده‌اند. یک روش ساده برای رمز کردن یک پیام m با استفاده از یک رمز قالبی با طول قالب n به صورت زیر است. اگر طول پیام مضرب n نباشد، ابتدا به روشی مناسب که دنباله‌زنی[۱۸] نامیده می‌شود، طول پیام به مضربی از n افزایش می‌یابد. سپس پیام دنباله‌زده شده به قالب های n-بیتی تقسیم می‌شود و هر قالب به طور جداگانه با اعمال رمزقالبی رمز می‌شود. برای رمزگشایی، نیز هر قالب متن رمزشده با اعمال معکوس رمز قالبی رمزگشایی می‌شود. یک روش مناسب و مرسوم دنباله‌زنی اضافه کردن یک بیت یک و سپس به تعداد لازم بیت صفر در انتهای پیام برای کامل کردن آخرین قالب است. دنباله اضافه شده پس از رمزگشایی قابل تشخیص و حذف است.



[۱۷] block cipher
[۱۸] padding



سؤال ۱۲ *آیا می‌توان برای دنباله‌زنی بیت یک را حذف کرد و فقط به تعداد لازم بیت صفر اضافه نمود؟*

سؤال ۱۳ *اگر آخرین قالب متن رمزگشایی شده تمام‌صفر باشد به چه معنایی است؟*

سؤال ۱۴ *این روش رمزنگاری با استفاده از رمز قالبی دارای کدام نوع از انواع امنیت است؟*