به نام خدا

دانشکده‌ی علوم ریاضی

مقدمه‌ای بر رمزنگاری ۱۳ اسفند ۱۳۹۱

جلسه‌ی هشتم: ویژگی خطی LFSRو رمزهای دنباله‌ای کلاسیک

مدرس: دکتر شهرام خزائی *نگارنده: سوزان اصغری*

۱ یادآوری

تعریف ۱ *(مرتبه‌ی چندجمله‌ای) برای چندجمله‌ای $P(x)$ که، $P(۰) \neq ۰$، کوچکترین عدد صحیح e که چندجمله‌ای $P(x)$، $x^e + ۱$ را عاد کند، مرتبه‌ی $P(x)$ است.*

تعریف ۲ *(چندجمله‌ای اولیه) چندجمله‌ای با درجه n (روی میدان متناهی $GF(۲)$) چندجمله‌ای اولیه است، هرگاه مرتبه‌ی آن $۲^n - ۱$ باشد.*

مثال ۳ *چندجمله‌ای $۱ + x + x^۴$ اولیه است، زیرا اگر متوالیا $۱ + x + x^۴$ را بر ۱ تقسیم کنیم، اولین باری که باقیمانده به‌صورت x^e باشد، برای $e = ۱۵$ اتفاق می‌افتد.*

$$
\begin{array}{r|l}
 & ۱ + x + x^۴ \\
\hline
۱ & \\
۱ + x + x^۴ & \\
\hline
x + x^۴ & ۱ + x + x^۲ + x^۳ + x^۵ + x^۷ + x^{۱۱} + x^۸ \\
- x + x^۲ + x^۵ & \\
\hline
x^۲ + x^۴ + x^۵ & \\
\vdots & \\
\hline
x^{۱۵} &
\end{array}
$$

البته الگوریتم‌های سریع‌تری برای آزمودن اولیه بودن یک چندجمله‌ای وجود دارد.

همان‌طور که در جلسه پیش دیدیم، دنباله خروجی LFSRها، مشروط بر اینکه چندجمله‌ای بازخورد آنها اولیه انتخاب شود، دارای دوره تناوب بالا ویژگی‌های آماری خوبی است. اما به خاطر ویژگی خطی بودن LFSRها، نباید آنها را مستقیما به عنوان یک مولد شبه‌تصادفی مورد استفاده قرار داد.

لم ۱ *برای هر $LFSR$ به طول L و با حالت اولیه $(s_۰, s_۱, ..., s_{L-۱})$، به ازای $t \geq ۰$، ضرایب $a_۰^t, a_۱^t, ..., a_{L-۱}^t$ وجود دارند، به طوری که:*

$$s_t = \sum_{i=۰}^{L-۱} a_i^t s_i \ .$$

برهان. فرض کنید $C(x) = ۱ - \sum_{i=۱}^{L} c_i X^i$ چندجمله‌ای بازخورد LFSRو $S_۰ = (s_۰, s_۱, ..., s_{L-۱})$ بردار حالت اولیه آن باشد. ماتریس انتقال حالت LFSRرا به‌صورت زیر تعریف می‌کنیم:

$$C = \begin{bmatrix} ۰ & ۰ & ۰ & \cdots & ۰ & c_L \\ ۱ & ۰ & ۰ & \cdots & ۰ & c_{L-۱} \\ ۰ & ۱ & ۰ & \cdots & ۰ & c_{L-۲} \\ \vdots & \vdots & \vdots & \cdots & \cdots & \\ ۰ & ۰ & ۰ & \cdots & ۱ & c_۱ \end{bmatrix} .$$

اگر بردار حالت LFSRرا در زمان $t \geq ۰$ با $S_t = (s_t, s_{t+۱}, ..., s_{t+L-۱})$ نشان دهیم، داریم $S_{t+۱} = S_t C$ و لذا برای بیت tام دنباله خروجی، s_t، داریم:

$$s_t = S_۰ C^t \begin{bmatrix} ۱ \\ ۰ \\ \vdots \\ ۰ \end{bmatrix} .$$

لذا ضرایب $a_۰^t, a_۱^t, ..., a_{L-۱}^t$ از رابطه زیر محاسبه می‌شوند:

$$\begin{bmatrix} a_۰^t \\ a_۱^t \\ \vdots \\ a_{L-۱}^t \end{bmatrix} = C^t \begin{bmatrix} ۱ \\ ۰ \\ \vdots \\ ۰ \end{bmatrix} .$$

■

مثال ۴ *برای $LFSR$ به طول ۴ و با چندجمله‌ای بازخورد $۱ + x + x^۴$ داریم $s_۴ = s_۱ + s_۰$، $s_۵ = s_۲ + s_۱$، $s_۶ = s_۲ + s_۳$ و به طور کلی*

$$s_t = [s_۰, s_۱, s_۲, s_۳] \begin{bmatrix} ۰ & ۰ & ۰ & ۰ & ۱ \\ ۱ & ۰ & ۰ & ۰ & ۱ \\ ۰ & ۱ & ۰ & ۰ & ۰ \\ ۰ & ۰ & ۰ & ۱ & ۰ \end{bmatrix}^t \begin{bmatrix} ۱ \\ ۰ \\ \\ ۰ \end{bmatrix} .$$

نتیجه ۵ *اگر L بیت $s_{t_۱}, s_{t_۲}, \cdots, s_{t_L}$ از دنباله خروجی یک $LFSR$ به طول L را بدانیم و ماتریس بردار ضرایب وارون‌پذیر باشد،حالت اولیه $LFSR$ به صورت زیر قابل محاسبه است*

$$S_۰ = (s_۰, s_۱, \cdots, c_{L-۱}) = (s_{t_۱}, s_{t_۲}, \cdots, s_{t_L}) \begin{bmatrix} a_۰^{t_۱} & a_۰^{t_۲} & \cdots & a_۰^{t_L} \\ a_۱^{t_۱} & a_۱^{t_۲} & \cdots & a_۱^{t_L} \\ \vdots & \vdots & \cdots & \vdots \\ a_{L-۱}^{t_۱} & a_{L-۱}^{t_۲} & \cdots & a_{L-۱}^{t_L} \end{bmatrix}^{-۱} .$$

نکته ۱ *احتمال اینکه یک ماتریس تصادفی روی $\mathbb{F}_۲$ معکوس‌پذیر باشد حداقل برابر $۰/۲۸۸ \approx \Pi_{i=۱}^{\infty}(۱ - \frac{۱}{۲^i})$ است.*

ممکن است به‌نظر برسد که با مخفی نگه داشتن چندجمله‌ای بازخورد LFSRبتوان مولد قویتری ساخت. اما با استفاده از مفهوم پیچیدگی خطی[1] که در ادامه معرفی می‌شود می‌توان نشان داد که مخفی نگه داشتن چندجمله‌ای بازخورد سطح امنیت را خیلی بالا نمی‌برد.

۲ پیچیدگی خطی

تعریف ۶ (*پیچیدگی خطی دنباله*) *پیچیدگی خطی دنباله* $(s_1, s_2, ..., s_n)$ *که با* $\Lambda(s_1, s_2, ..., s_n)$ *نشان داده می‌شود، طول کوچکترین* $LFSR$ *ی است که دنباله را تولید می‌کند.*

پیچیدگی خطی یک دنباله طول کوچکترین LFSRای است که آنرا تولید می‌کند. اگر دنباله داده شده، دنباله خروجی یک LFSRبه طول L با چندجمله‌ای بازخورد اولیه و از یک حالت اولیه ناصفر باشد، آنگاه پیچیدگی خطی آن دقیقا L خواهد بود.
با استفاده از الگوریتم Berlekamp-Massey می‌توان حالت اولیه و چندجمله‌ای بازخورد کوچکترین LFSRای که یک رشته را تولید می‌کند محاسبه کرد.

قضیه ۲ *اگر پیچیدگی خطی دنباله‌ای* Λ *باشد، چندجمله‌ای بازخورد و حالت اولیه کوچکترین* $LFSR$ *ی که دنباله را تولید می‌کند، از روی هر* 2Λ *بیت متوالی به صورت یکتا و با پیچیدگی* $O(\Lambda^2)$ *محاسبه می‌شود.*

۳ رمزهای دنباله‌ای مبتنی بر LFSR

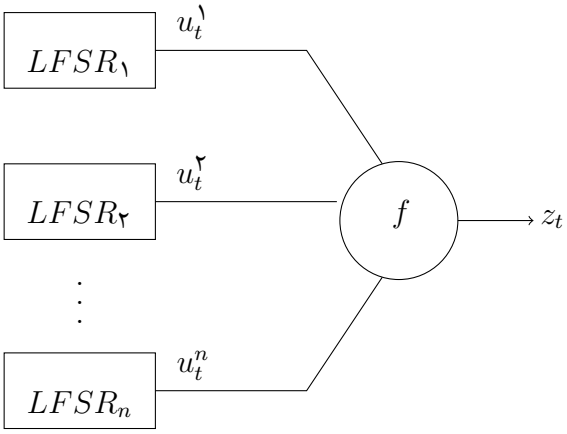
برای از بین بردن ویژگی خطی LFSRها باید از یک عامل غیر خطی در تولید دنباله خروجی مولد (که دنباله کلید اجرایی[2] نامیده می‌شود) استفاده کرد. برای این منظور سه روش کلی وجود دارد که در ادامه معرفی می‌شوند.

۱.۳ مولدهای ترکیبی

مولد ترکیبی[3] مولدی است که دنباله‌ی کلید اجرایی آن از ترکیب دنباله‌های خروجی چند LFSRبا استفاده از یک تابع بولی غیر خطی f بدست می‌آید. اگر دنباله‌های خروجی n عدد LFSR، که با u_t^i نشان داده می‌شوند، با استفاده از تابع بولی n-متغیره f با هم ترکیب شوند، tامین بیت دنباله کلید اجرایی، z_t، از رابطه زیر محاسبه می‌شود:

$$z_t = f(u_t^1, u_t^2, ..., u_t^n)$$

[1] linear complexity
[2] running-key or keystream
[3] combination generator



برای طراحی یک مولد ترکیبی که دنباله آن شبه تصادفی باشد باید طول LFSRها و تابع ترکیب‌کننده f به دقت انتخاب شوند.

قضیه ۳ *اگر طول همه‌ی $LFSR$ها که با $L_1, L_2, .., L_n$ نمایش می‌دهیم، متمایز و بزرگتر از ۲ باشند و نیز حالت‌های اولیه ناصفر باشند، پیچیدگی خطی دنباله‌ی خروجی عبارت است از:*

$$f(L_1, L_2, .., L_n)$$

معیارهای لازم برای انتخاب تابع بولی f:

- متعادل[۴] باشد.
- درجه‌ی جبری[۵] بالایی داشته باشد.
- مرتبه امنیت همبستگی[۶] بالایی داشته باشد.
- مقدار غیر خطی[۷] بالایی داشته باشد.
- مرتبه امنیت جبری[۸] بالایی داشته باشد.

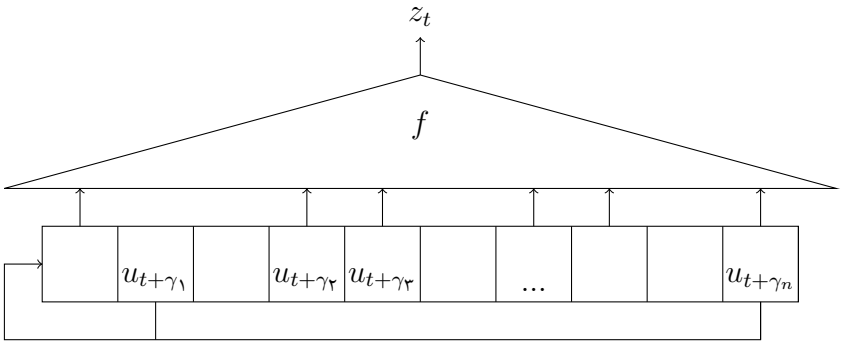
همچنین جهت دشوارتر کردن اعمال حمله همبستگی[۹]، وزن‌های چندجمله‌ای‌های فیدبک باید بزرگ باشند.

[۴] balanced
[۵] algebraic degree
[۶] correlation-immunity order
[۷] non-linear order
[۸] algebraic-immunity order
[۹] correlation attack

۲.۳ مولد فیلتر غیرخطی

مولد فیلتر غیرخطی[۱۰] بجای ترکیب دنباله خروجی چند LFSRمختلف، انتقال یافته‌های دنباله‌ی خروجی یک عدد LFSRرا با هم ترکیب می‌کند. برای سادگی پیاده‌سازی، خروجی این مولد از اعمال یک فیلتر غیر خطی n-متغیره بر n سلول از سلول‌های یک LFSRبه‌طول L (که $n \leq L$) که از موقعیت‌های $0 \leq \gamma_1 < \gamma_2 < \gamma_n \leq L-1$ گرفته‌می‌شوند به دست می‌آید. دقت کنید که دنباله‌ای که از سلول γ_iام LFSRگرفته می‌شود، انتقال‌یافته دنباله خروجی LFSRبه اندازه γ_i واحد زمانی است. به عبارت دقیق‌تر اگر u_t دنباله خروجی LFSRباشد ($t \geq 0$)، tامین بیت دنباله کلید اجرایی، z_t، از رابطه زیر محاسبه می‌شود:

$$z_t = f(u_{t+\gamma_1}, u_{t+\gamma_2}, ..., u_{t+\gamma_n}) .$$



برای اینکه دنباله کلید اجرایی شبه تصادفی باشد، چندجمله‌ای فیدبک LFSRباید چندجمله‌ای اولیه باشد و تابع فیلتر f ویژگی‌هایی را که برای مولد ترکیبی گفته شد داشته باشد. همچنین وزن چندجمله‌ای فیدبک باید بالا باشد و موقعیت سلول‌ها باید به درستی انتخاب شود تا از امکان اعمال بعضی از حملات جلوگیری شود.
دنباله‌ی خروجی مولد فیلتر یک دنباله‌ی دوره‌ای است که پیچیدگی خطی آن به طول LFSRو درجه جبری تابع وابسته‌است.

قضیه ۴ *برای یک $LFSR$ باینری به طول L و درجه‌ی جبری d که چندجمله‌ای فیدبک آن چندجمله‌ای اولیه باشد، عبارت زیر برقراراست:*

$$\Lambda \leq \sum_{k=0}^{d} \binom{L}{k}$$

همچنین اگر L عدد بزرگی باشد، برای اکثر توابع فیلتر با درجه‌ی جبری d، کران پایین $\Lambda \leq \binom{L}{d}$ صادق است.

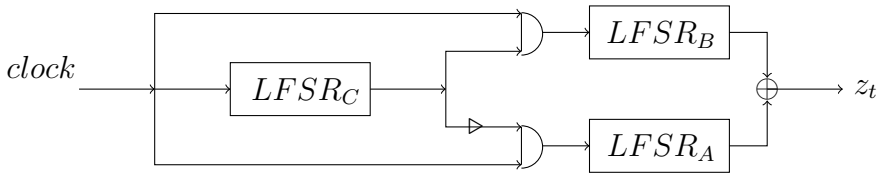
۳.۳ مولدهای با انتقال‌های نامنظم

یک روش دیگر برای وارد کردن عامل غیر خطی بدون استفاده از فیلتر غیرخطی، استفاده از LFSRهای با انتقال نامنظم[۱۱] است. سه مولد کلاسیک زیر از این ایده استفاده می‌کنند که در عین سادگی هنوز حمله کارایی علیه آنها پیدا نشده است.

[۱۰] nonlinear filter generator
[۱۱] clock-controlled LFSR

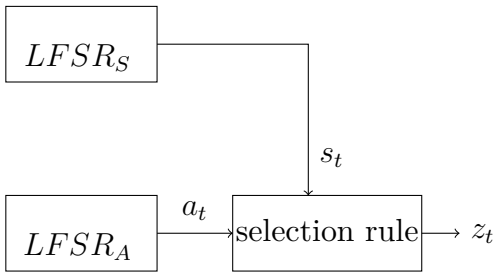
- مولد گام متناوب [۱۲]
- مولد کاهنده [۱۳]
- مولد خودکاهنده [۱۴]

مولد گام متناوب این مولد از سه LFSRتشکیل شده‌است: $LFSR_A$، $LFSR_B$ و $LFSR_C$ که $LFSR_C$ کنترل کننده است دو LFSRدیگر است. $LFSR_C$ به‌طور منظم انتقال می‌یابد و خروجی آن تعیین می‌کند که کدام یک از دو LFSRدیگر انتقال یابد: اگر خروجی $LFSR_C$ یک باشد $LFSR_A$ انتقال می‌یابد و اگر صفر باشد $LFSR_B$ کلاک می‌خورد. پس از هر انتقال، یک دنباله کلید اجرایی تولید می‌شود که از XOR (جمع مبنای دو) بیت‌های خروجی LFSRهای کنترل‌شونده بدست می‌آید.



نسخه دیگری از مولد گام متناوب وجود دارد که در آن خروجی به‌جای XOR کردن بیت‌های خروجی LFSRهای کنترل‌شونده، از همان LFSRای که انتقال یافته است بدست می‌آید. به عبارت دیگر دنباله کلید اجرایی، از لای هم قرار دادن[۱۵] خروجی‌های منظم LFSRهای کنترل‌شونده حاصل می‌شود. می‌توان نشان داد که امنیت این دو نسخه از مولد گام متناوب یکسان است.
برای داشتن ویژگی بهتر دنباله خروجی بهتر است طول LFSRها دوبه‌دو نسبت به هم اول باشند که در اینصورت، دنباله خروجی مولد به طرز قابل اثباتی دارای دوره‌ی بلند و پیچیدگی خطی بالاست. اگر طول هر سه LFSRمورد استفاده تقریبا برابر L باشد، پیچیدگی بهترین حمله‌ی شناخته شده علیه این مولد از مرتبه $O(۲^{۲L/۳})$ است.

مولد کاهنده این مولد از دو LFSRتشکیل شده‌است: $(LFSR_S, LFSR_A)$. دنباله کلید اجرایی از $LFSR_A$ گرفته می‌شود و خروجی $LFSR_S$ تعیین می‌کند که کدام بیت‌های $LFSR_A$ به عنوان خروجی درنظر گرفته شوند.
مولد کاهنده به این صورت عمل می‌کند: اگر بیت خروجی $LFSR_S$ یک باشد، بیت خروجی $LFSR_A$ به عنوان یک بیت از دنباله کلید اجرایی مولد به خروجی فرستاده می‌شود و اگر صفر باشد، بیت خروجی $LFSR_A$ دور انداخته می‌شود یعنی هیچ بیتی تولید نمی‌شود؛ سپس هر دو LFSRانتقال می‌یابند.



[۱۲] Alternating Step Generator(ASG)
[۱۳] Shrinking Generator(SG)
[۱۴] Self-Shrinking Generator(SSG)
[۱۵] interleaving

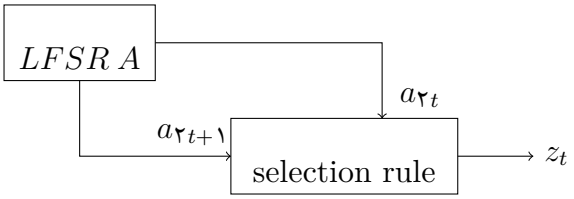
به عبارت دیگر اگر دوبیت (s_t, a_t) برابر ۱۰ باشد، خروجی ۰ است، اگر خروجی ۱۱ باشد، خروجی ۱ است. در غیر اینصورت خروجی نداریم.
اگر طول هر دو LFSRمورد استفاده تقریبا برابر L باشد، پیچیدگی بهترین حمله‌ی شناخته شده علیه این مولد از مرتبه $O(۲^{۲L/۳})$ است.

مولد خودکاهنده اساس این مولد همان مولد قبلی است که در آن از یک LFSRبا دنباله خروجی $\{a_t\}$ استفاده می‌شود که در آن زیردنباله $\{a_{۲t}\}$ نقش خروجی $LFSR_S$ و زیردنباله $\{a_{۲t+۱}\}$ نقش خروجی $LFSR_A$ را دارد. به طور دقیق‌تر رویه زیر نحوه کار مولد خودکاهنده را نشان می‌دهد.

تا وقتی که دنباله کلید اجرایی به طول کافی تولید نشده است:

۱. LFSRدوبار کلاک می‌خورد و دو بیت $(a_{۲t}, a_{۲t+۱})$ از آن گرفته می‌شود.

۲. اگر $a_{۲t} = ۱$ باشد، $a_{۲t+۱}$ به عنوان یک بیت از دنباله خروجی تولید می‌شود.



به عبارت دیگر اگر دوبیت خروجی ۱۰ باشد، خروجی ۰ است، اگر خروجی ۱۱ باشد، خروجی ۱ است. در غیر اینصورت خروجی نداریم.
بهترین حمله‌ی شناخته شده علیه این مولد حمله بده‌بستان حافظه-زمان-داده است که به حافظه، داده (طول دنباله خروجی مورد نیاز) و زمان از مرتبه $O(۲^{L/۲})$ نیاز دارد. پیدا کردن یک حمله بهتر یک پرسش باز جالب است.