پیرامون ANF و درجه جبری عمل دوران وابسته به داده

اکبر شاهسواران[۱]

چکیده

در این مقاله عمل دوران وابسته به داده(DDR) را به عنوان یک تابع بولی برداری در نظر می گیریم و نشان می دهیم که درجه جبری همه توابع مولفه ای آن k+۱ است که 2^K طول بیتی عملوندهاست.این موضوع به علاوه فرم نرمال جبری (ANF) توابع مولفه ای که ارایه می شود، علاوه بر اهمیت نظریهک و بینشی که نسبت به ماهیت جبری این عمل به ظاهر رام نشدنی و پیچیده به دست میدهد، کاربردهایی در بررسی حمله تفاضلهای مراتب بالا بر الگوریتمهای قالبی نظیر RC۵ که از DDR به عنوان یک مولفه استفاده می کنند و نیز در بررسی حمله جبری Courtois دارد.

کلید واژه

دوران وابسته به داده، فرم نرمال جبری، درجه جبری، حمله تفاضلهای مراتب بالا ، حمله جبری Courtois

۱. دانشگاه صنعتی شریف a_shahsavaran@yahoo.com

تاریخ دریافت: ۸۸/۱۲/۱۰ تاریخ پذیرش: ۸۹/۱/۱۵

مقدمه

اگر x وy دو 2^K بیتی باشند، آنگاه منظور از x<<<y دوران یافته x به چپ به اندازه مقدار دهدهی k بیت با ارزش کمتر از y است. هنگامی که در یک الگوریتم رمز ، y کمیتی برگرفته از حالت (State) داخلی الگوریتم باشد، به این عمل ، اصطلاحاً عمل دوران وابسته به داده یا Data-Dependent Rotation گفته می‌شود. این عمل نخستین بار به طور گسترده در الگوریتم رمز RC5 به کار گرفته شد [۳] و بعدا در الگوریتمهای دیگری نظیر RC6 [۶] و MARS [۷] به کار گرفته شد. خواص خطی و تفاضلی این عمل به ترتیب در [۵] و [۸] مورد بررسی قرار گرفته است. ما در این نوشته درجه جبری این عمل را مورد بررسی قرار می‌دهیم . دانستن درجه جبری یک تابع علاوه بر کمک به ارزیابی قدرت آن در مقابله با حمله تفاضلهای مراتب بالاتر[۲]، از نظر حمله جبری Courtois [۴] نیز مهم است. سازماندهی ادامه مقاله به این صورت است: در بخش ۲ درجه جبری و فرم نرمال جبری عمل دوران وابسته به داده را محاسبه می کنیم و ساده سازی هایی را که در محاسبه فرم نرمال جبری می توان انجام داد، می آوریم. در بخش ۳ نتایج و کاربردهای موضوع مقاله را بر می شماریم. در بخش ۴ به موضوعات مرتبطی که جای کار دارند، اشاره می کنیم. مراجع در پایان آورده خواهند شد.

موضوعات و روشها

درجه جبری عمل DDR

قضیه : به ازای هر n طبیعی، هر تابع بولی $f : Z_2^n \to Z_2$ دارای نمایشی یکتا به صورت زیر است :

$$f(x) = \bigoplus_{(u_1,..,u_n)\in Z_2^n} h_{u_1,..,u_n} x_1^{u_1} ... x_n^{u_n}$$

که $h_{u_1,..,u_n} \in Z_2$ برای هر $(u_1,..,u_n) \in Z_2^n$.

اثبات: [۱]

تعریف : نمایش یکتای مذکور در قضیه فوق را فرم نرمال جبری (ANF) تابع بولی f می‌نامند. به تعداد u_i های ناصفر در عبارت $x_1^{u_1} ... x_n^{u_n}$ درجه f نسبت به جمله $x_1^{u_1} ... x_n^{u_n}$ ، و به بزرگترین این درجات، درجه جبری f اطلاق می شود. بدیهی است همواره درجه جبری f نابیشتر از n است.

اکنون عمل x<<<y را که قبلاً تعریف شد، در نظر بگیرید. فرض می‌کنیم بیت‌های سمت راست، بیت‌های باارزش کمتر باشند؛ پس مقدار x<<<y تنها به x و k بیت سمت راست y وابسته است.

اگر $x = x_{2^k-1}....x_0$ و $y = y_{2^k-1}y_{2^k-2}...y_0$ آنگاه می‌توان نوشت:

$$x <<< y = F(x_{2^k-1},...,x_0, y_{k-1}, y_{k-2},...y_0) \qquad (1)$$

و F تابعی است که $2^k + k$ بیت را به 2^k بیت تصویر می‌کند. فرض کنید $F = (f_{2^k-1},...,f_0)$، یعنی f_i ها توابع مؤلفه‌ای F باشند. نیز فرض کنید $F(x_{2^k-1},...,x_0, y_{k-1},...,y_0) = (z_{2^k-1},...,z_0)$. در این صورت برای هر $i = 0,...,2^k - 1$:

$$f_i(x_{2^k-1},...,x_0, y_{k-1},...,y_0) = z_i$$

برای به دست آوردن ANF تابع f_i ، از ویژگی یکتایی ANF که در قضیه فوق ذکر شد، استفاده می‌کنیم . در واقع z_i یکی از x_0 ، x_1 ، ... یا x_{2^k-1} است و این بستگی به مقدار y_0 ،.. y_{k-1} و دارد. پس برای هر i می توان نوشت : (*)

$$f_i(x_{2^k-1},...,x_0, y_{k-1},...,y_0) = x_0 f_{i,0}(y_{k-1},...,y_0) \oplus ... \oplus x_{2^k-1} f_{i,2^k-1}(y_{k-1},...,y_0)$$

که $f_{i,j}$ ها توابعی هستند که باید به دست بیایند. از همین فرمول به دست می آید که درجه جبری f_i حداکثر k+۱ است. نشان خواهیم داد که این عدد دقیقا k+۱ است.

ANF و درجه توابع $f_{i,j}$

از رابطه (۱) و (*) به دست می‌آید که :

$$y = 0 \rightarrow f_{i,i} = 1 \quad , \quad f_{i,j} = 0, \quad \forall j \neq i$$
$$y = 1 \rightarrow f_{i,i-1} = 1 \quad , \quad f_{i,j} = 0, \quad \forall j \neq i-1$$
$$\vdots$$
$$y = 2^k - 1 \rightarrow f_{i,i-(2^k-1)} = 1 \quad , \quad f_{i,j} = 0 \quad \forall j \neq i-(2^k-1)$$

که در اینجا $y = (y_{k-1},...,y_0)$ و محاسبات اندیس‌ها به پیمانه 2^k انجام می‌شود.

اکنون برای یک i ثابت ، و برای یک j ثابت، اگر $j = (i-m) \bmod 2^k$ ، آنگاه از اطلاعات سطر y=m به دست می‌آید $f_{i,j}(m) = 1$ و از هر سطر دیگر $y = m'(\neq m)$ به دست می‌آید که $f_{i,j}(m') = 0$ زیرا $j \neq i - m'$.

$$f_{i,i-m}(y)=\begin{cases}1; y=m\\ 0; y\neq m\end{cases} \qquad \text{پس} \qquad (2)$$

اکنون چون مقادیر تابع $f_{i,j}$ را برای هر مقدار متغیر در اختیار داریم، می‌توانیم فرم نرمال جبری آن را به دست آوریم . برای محاسبه ANF توابع $f_{i,j}$ ، ابتدا دسته چندجمله ایهای g_m را تعریف می کنیم.

تعریف : به ازای هر k بیتی $m=(m_{k-1},...,m_0)$ تابع چندجمله ای $g_m : Z_2^k \to Z_2$ را به صورت زیر تعریف می کنیم:

$$g_m(y_{k-1},...,y_0)=(y_{k-1}\oplus m_{k-1}\oplus 1)(y_{k-2}\oplus m_{k-2}\oplus 1)...(y_0\oplus m_0\oplus 1)$$

به وضوح $g_m(y)=1$ اگر و تنها اگر $y=(m_{k-1},...,m_0)=m$.

ملاحظه می شود که برای هر y : $f_{i,i-m}(y)=g_m(y)$ و به دلیل یکتایی ANF توابع بولی که در قضیه ابتدای مقاله به آن اشاره شد، فرم نرمال جبری $f_{i,i-m}(y)$ به صورت زیر است:

$$f_{i,i-m}(y_{k-1},...,y_0)=(y_{k-1}\oplus m_{k-1}\oplus 1)(y_{k-2}\oplus m_{k-2}\oplus 1)...(y_0\oplus m_0\oplus 1) \quad (**)$$

نتیجه: بلافاصله به دست می آید که درجه جبری $f_{i,i-m}$ مساوی k است.

بنابراین برای هر i و j ، برای به دست آوردن $ANF(f_{i,j})$ ، k بیتی m را چنان می یابیم که $j=(i-m)\bmod 2^k$ ، و از رابطه (**) استفاده می کنیم.

ANF و درجه توابع f_i

قضیه اصلی : درجه جبری هر تابع f_i دقیقا مساوی $k+1$ است.

اثبات: طبق نتیجه ای که به دست آوردیم، درجه جبری هر تابع $f_{i,j}$ مساوی k است و بنابراین درجه هر تابع $x_j f_{i,j}(y_{k-1},...,y_0)$، $k+1$ است. بدیهی است که در حاصل‌جمع (*) بین جملات حذف صورت نگرفته و درجه جبری $k+1$ خواهد ماند.

رابطه (*) به علاوه نشان می دهد که درجه f_i نسبت به هر کدام از متغیرهای x_j مساوی ۱ است.

با جایگذاری فرم‌های نرمال جبری $f_{i,j}$ ها در (*)، فرم نرمال جبری f_i به دست می‌آید.

ارتباط ANF توابع $f_{i,j}$

رابطه (*)شکل کلی فرم نرمال جبری هر کدام از توابع f_i را در اختیار قرار می دهد. تنها کافی است ANF توابع $f_{i,j}$ که تعداد متغیرهای ورودی آنها k است، تعیین شود. از رابطه (۲) به دست می آید که اگر $i-j$ و $i'-j'$ به پیمانه 2^k همنهشت باشند، آنگاه توابع $f_{i,j}$ و $f_{i',j'}$ کاملا مساوی خواهند بود. بنابراین مثلا کافی است ANF توابع $f_{1,0}$ ، $f_{1,1}$ ،...و $f_{1,2^k-1}$ را محاسبه کنیم. در میان این توابع نیز کافی است مثلا برای $f_{1,1}$ محاسبه ANF صورت گیرد زیرا از (۲) به دست می آید که $f_{1,0}(x)=f_{1,1}(x\oplus 1)$ و برای $j\geq 2$ ، $f_{1,j}(x)=f_{1,1}(x\oplus(2^k+1-j))$.

نتایج

الف- در نگاه اول کاربرد آنی نتایج این مقاله در ارتباط با حملاتی است که از پایین بودن درجه جبری عملهای رمزنگاری بهره می گیرند، نظیرحمله تفاضلهای مراتب بالاتر در رمزهای قالبی.

ب-کاربرد دیگر که به نظر ما بسیار مهم است، در ارتباط با حملات جبری ارائه شده توسط Courtois است [مثلا ۱ و۲] .

ب_۱ در حمله جبری به رمزهای دنباله ای حمله کننده در صدد است رابطه ای مستقیم بین بیتهای دنباله کلید اجرایی و بیتهای حالت داخلی الگوریتم به دست آورد و با تشکیل یک دستگاه معادلات بر حسب بیتهای حالت اولیه، آن را حل کند. صرف نظر از روشهای حل دستگاه، در مرحله اول حمله، هدف تشکیل یک دستگاه معادلات است . آنچه در این مقاله به دست آوردیم، از جمله ANF توابع f_i و درجه آنها، و ویژگیهای توابع $f_{i,j}$ که امکان دست ورزیهای جبری(Algebraic Manipulations) را ممکن است فراهم کنند، می تواند در راستای تحلیل جبری الگوریتمهایی که از عمل دوران وابسته به داده استفاده می کنند، کمک نماید.

ب_۲ در رمزهای قالبی نیز ANF توابع f_i و درجه نسبتا پایین آنها یک سری روابط حداقلی را بین ورودی و خروجی تابع دوران (2^k+k متغیر ورودی و 2^k متغیر خروجی) در اختیار حمله کننده قرار می دهند و این ذهنیت را که عمل دوران وابسته به دیتا (در رمزهایی نظیر RC6 و MRAS) یا دوران وابسته به کلید(در رمزهایی نظیر خانواده CAST) به طور کلی مانع انجام حمله جبری به اینگونه رمزها می شود، از بین می برد(لازم به ذکر است برای اعمال دیگری نظیر جمع یا تفاضل پیمانه ای که در این الگوریتمها به کار رفته، نیز می توان مجموعه ای از معادلات درجه ۲ نوشت که آنها را کاملا توصیف کند).

توضیح آنکه اگر k کوچک باشد، در بحث حمله جبری می توان تابع دوران وابسته به داده:

$$F(x_{2^k-1},...,x_0,y_{k-1},...,y_0)=(z_{2^k-1},...,z_0)$$

را یک جعبه جانشینی با ابعاد $(2^k+k)\times 2^k$ بیت دانست و با روشی که در منابع گفته شده، به دنبال معادلات ضمنی بین بیتهای ورودی و بیتهای خروجی گشت؛ ولی هنگامی که k بزرگ باشد، این کار عملی نیست و روابط درجه پایینی که ما برای z ها بر حسب x ها و y ها به دست آورده ایم به کار گرفته خواهند شد. به عنوان یک مساله پژوهشی جالب می توان در ادامه و تکمیل کار مقاله حاضر، مساله یافتن سامانه اتیک معادلات ضمنی درجه پایین یا overdefined یا با ویژگیهای دیگری که به حمله جبری کمک کند، را مطرح نمود، به گونه ای که از آنچه در این مقاله به دست آوردیم، بهتر عمل کنند.

پ- در اختیار داشتن فرمولهای ریاضی و جبری برای توصیف اعمالی که اصالتا به گونه ای دیگر تعریف شده اند نظیر دوران وابسته به داده، این امکان را به وجود می آورد که در صورت نیاز به انجام پژوهش بر روی ویژگیهای این اعمال، از این فرمولها و خواص اثبات شده استفاده کرد.

ت- بدیهی است که گزاره ها و نتایجی که در مورد سرشت جبری عمل دوران وابسته به داده در این مقاله به دست آوردیم، برای حالتی که بیتهای تعیین کننده مقدار دوران از بیتهای با ارزش بیشتر یا هر مکان دیگر از y انتخاب شوند، نیز صادق است. به علاوه به سادگی می توان دید که حتی برای حالتی که اصطلاحا عمل خود-دوران داریم و به جای k ، y بیت از خود x مقدار دوران x را تعیین می کنند، نیز نتایج مشابهی داریم.

بحث

عمل دوران وابسته به داده یا دوران وابسته به کلید، به شیوه های غیر جبری توصیف شده اند. در این مقاله شکل نرمال جبری آنها را ارائه کردیم و مقدار دقیق درجه جبری توابع مولفه ای را به دست آوردیم. در اختیار داشتن این ویژگیهای جبری برای عملی که به صورت غیر جبری توصیف می شود و در ظاهر هیچ ویژگی جبری مشخصی برای آن نمی توان برشمرد، بینش ما را نسبت به سرشت رمزنگاری این عمل زیاد می کند و به علاوه کاربردهایی دارد که در بخش قبل برشمرده شدند. از کارهای پژوهشی مناسبی که لازم است در تکمیل این مقاله انجام شود، بررسی نوع معادلات ضمنی (معادلات صریح ، در این مقاله به دست آمدند) است که می توان در ارتباط با حمله جبری Courtois بین بیتهای ورودی و خروجی به دست آورد و ما در آینده روی این موضوع کار خواهیم کرد.

مراجع

1.Steven Roman, “Coding and information theory” , Springer-Verlag, 1992

2.L.R. Knudsen, “Truncated and Higher Order Differentials” FSE, 2nd Intl Workshop Proceedings,Springer, 1995, pp. 196

3.R.L.Rivest,”The RC5 Encryption Algorithm”FSE, LNCS 1008, pp. 86-96,Springer-Verlag, 1995

4.N. Courtois , J. Pieprzyk, "Cryptanalysis of block ciphers with overdefined systems of equations" in Proc. of Asiacrypt 2002 LNCS 2501, Springer-Verlag 2002

5.S. Moriai et al, "Key-Dependency of Linear Probability of RC5",IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E80-A, No.1, (Jan., 1997)

6.R.L. Rivest et al, “The RC6 Block Cipher”,v1.1, 1998, Available from www.rsasecurity.com/rsalabs/aes/

7.C.Burwick et al, “ MARS : a candidate cipher for AES” IBM Corporation , June 10 , 1998

8.S. Contini et al, “On Differential Properties of Data-Dependent Rotations and Their Use in MARS and RC6”, Presented at the 2nd AES Conference