مجله علمی-پژوهشی «علوم و فناوری‌های پدافند غیرعامل»
سال دوم، شماره ۲، تابستان ۱۳۹۰: ص ۱۰۰- ۹۱

تحلیل دو نمونه ساده شده تابع چکیده‌ساز MD4، با استفاده ازخطی‌سازی

جواد علی‌زاده*[۱]،جواد مهاجری[۲]، نصور باقری[۳]

۱- کارشناس ارشد، دانشگاه جامع امام حسین (ع)، مرکز تحقیقات فتح

۲- استادیار، دانشگاه صنعتی شریف، پژوهشکده الکترونیک

۳- استادیار، دانشگاه شهید رجایی، دانشکده مهندسی برق و کامپیوتر

(دریافت: ۱۳۸۹/۰۷/۲۰، پذیرش: ۱۳۹۰/۰۳/۲۳)

چکیده

امنیت اطلاعات، یکی از مقوله‌های مهم پدافند غیرعامل بوده و علم رمزنگاری در برقراری این امنیت، نقش انکارناپذیری دارد. برای این منظور، از توابع چکیده‌ساز رمزنگاری، با هدف اصلی ایجاب تمامیت در اطلاعات استفاده می‌شود. یکی از روش‌های مهم برای تحلیل توابع چکیده‌ساز رمزنگاری، استفاده از خطی‌سازی است که از آن برای تحلیل بسیاری از توابع درهم‌ساز استفاده شده است. در این مقاله، ایده روش تحلیل ذکر شده، همراه با جبر خطی، استفاده و با یک الگوریتم جستجوی جدید، چند الگوی تفاضلی برای تابع چکیده‌ساز MD4 خطی (LMD4) به‌دست آمده است. سپس با استفاده از یکی از این الگوها، یک برخورد برای LMD4 محاسبه شد. همچنین یک نوع ساده شده از تابع چکیده‌ساز MD4، به اسم تابع چکیده‌ساز CMD4، تعریف و با استفاده از یکی از الگوهای تفاضلی ذکرشده، یک حمله برخورد موفقیت‌آمیز با پیچیدگی 2^{35} روی آن اعمال شد. در نهایت الگوهای تفاضلی ذکر شده برای حمله برخورد روی خود تابع چکیده‌ساز MD4 بررسی شدند.

کلیدواژه‌ها: تابع چکیده‌ساز، MD4، حمله برخورد، خطی‌سازی، الگوی تفاضلی

Cryptanalysis of Two Simplified Variants of MD4, Using Linearization

J. Alizadeh*[1], J. Mohajeri[2], N. Bagheri[3]

Fath Research Center, Imam Hossein University

(Received:10/12/2010, Accepted:06/13/2011)

Abstract

Information security is one of the important aspects of passive defense and cryptography has an undeniable role in realization of this security. Therefore, the main aim of using cryptographic hash functions is the integrity of information. Linearization is one of the important techniques in cryptanalysis of the hash functions that used for cryptanalysis of many hash functions. In this paper, we use idea of the cryptanalysis method with linear algebra and find some differential patterns for a linear variant of MD4 (LMD4) with a new search algorithm. Then using one of the patterns we find a collision for LMD4. Also we define another simplified variant of MD4 (CMD4) and present a successful collision attack with a complexity of 2^{35} on the CMD4 by using one of the so-called patterns is achieved. Finally, the collision attack on the MD4 is reconsidered.

Keywords: Hash Function, MD4, Collision Attack, Linearization, Differential Pattern

* Corresponding author E-mail: alizadja@gmail.com

۱. مقدمه

ایجاب امنیت اطلاعات (در طول برقراری یک ارتباط یا بدون داشتن ارتباط)، یکی از مؤلفه‌های اساسی در پدافند غیرعامل است. در تعریف این امنیت، می‌توان محرمانگی[۱]، جامعیت[۲] و دسترس‌پذیری[۳] را به عنوان سه اصل اساسی آن و همچنین علم رمزنگاری[۴] را به عنوان یک ابزار مهم (و شاید مهم‌ترین ابزار)، برای برآوردن برخی از اصل‌های آن بیان کرد.

توابع چکیده‌ساز از اولیه‌های رمزنگاری[۵] مهم هستند که هدف اصلی استفاده از آنها، ایجاب جامعیت در امنیت اطلاعات می‌باشد، به طوری که می‌توان کاربرد آنها را در حوزه‌های وسیعی شامل امضاهای رقمی[۶] و طرح‌های احراز هویت[۷] مشاهده کرد. در بسیاری از موارد، امنیت طرح‌های مذکور ارتباط مستقیمی با امنیت تابع چکیده‌ساز به کار رفته در آنها دارد.

یک تابع چکیده‌ساز رمزنگاری، نگاشتی مانند $h:\{0,1\}^* \to \{0,1\}^n$ است که در آن n یک عدد صحیح مثبت است (منظور از $\{0,1\}^*$ و $\{0,1\}^n$، به ترتیب رشته‌بیت‌های به طول دلخواه و با اندازه ثابت n می‌باشد). به عبارت دیگر، یک تابع چکیده‌ساز، نگاشتی است که یک ورودی با طول متغیر را به یک خروجی با اندازه ثابت (که خلاصه پیام نامیده می‌شود) نگاشت می‌کند. تعریف دقیق‌تری از یک خانواده چکیده‌ساز[۸] را می‌توان مطابق تعریف (۱) بیان کرد:

تعریف ۱. (تعریف دقیق تابع چکیده‌ساز، با توجه به توزیع احتمال):
فرض کنیم $M=\{0,1\}^n$ فضای پیام‌های ورودی، $Y=\{0,1\}^m$ فضای خروجی و $K=\{0,1\}^t$ فضای کلیدهای ممکن باشد، آنگاه برای هر $x\in M$ و هر $k\in K$ حداقل یک $y\in Y$ وجود دارد به‌طوری که،

$$h:\{0,1\}^t\times\{0,1\}^n \to \{0,1\}^m$$
$$h_k(x)=h(k,x)=y.$$

در این صورت گوییم تابع h یک تابع چکیده‌ساز است. هرگاه دو شرط زیر برای یک تابع چکیده‌ساز برقرار باشند، تابع چکیده‌ساز، یک تابع چکیده‌ساز قوی نامیده می‌شود.

i) $\forall k\in K,\ x\in M,\ y\in Y \Rightarrow p[h_k(x)=y]=\frac{1}{2^m}$

ii) $\forall k\in K,\ \forall x,x^*\in M,\ \forall y,y^*\in Y \text{ if } x\neq x^*$
$\Rightarrow p[h_k(x)=y \,\&\, h_k(x^*)=y^*]=\frac{1}{2^{2m}}$

شرط اول بیان می‌کند که تابع چکیده‌ساز باید عناصر دامنه را با احتمال یکسان به عناصر برد تصویر کند و شرط دوم بیان می‌کند که زوج‌های دلخواه و تصادفی (x,x^*) می‌بایست مستقل از زوج‌های توابع چکیده‌ساز $(h_k(x),h_k(x^*))$ باشند. در عمل، رسیدن به چنین فرضی مشکل است، لذا بحث اصلی در این راستا، رسیدن به تابعی می‌باشد که به ازای هر ورودی دلخواه، خروجی تابع چکیده‌ساز به سمت تصادفی شدن متمایل گردد، بدین مفهوم که خروجی شبه تصادفی[۹] باشد. این یک ملاک معتبر و مناسب برای طراحی توابع چکیده‌ساز می‌باشد. در تعریف فوق اگر مقدار k ثابت باشد، تابع چکیده‌ساز را بدون کلید و در غیر این‌صورت، آن را تابع چکیده‌ساز کلیددار گوییم. (در این مقاله منظور از تابع چکیده‌ساز، همان تابع چکیده‌ساز بدون کلید فرض می‌شود).

یک تابع چکیده‌ساز رمزنگاری امن، می‌بایست سه ویژگی امنیتی مقاومت در برابر پیش‌تصویر[۱۰]، مقاومت در برابر پیش‌تصویر دوم[۱۱] و مقاومت در برابر برخورد[۱۲] را ایجاب کند. بدین معنی که برای یک تابع چکیده‌ساز امن با داشتن یک مقدار چکیده‌ساز مانند v، یافتن پیامی مانند x، به‌طوریکه h(x)=v باشد (پیش‌تصویر)، یا با داشتن پیامی مانند y، یافتن پیامی مانند $x\neq y$، به‌طوری که h(x)=h(y) باشد (پیش‌تصویر دوم) و یا یافتن دو پیام مختلف x و x' به‌طوری که $h(x)=h(x')$ باشد(برخورد)، از لحاظ محاسباتی غیرعملی باشد. غیرعملی بودن از لحاظ محاسباتی، به این معنی است که با استفاده از امکانات محاسباتی امروزی، نتوان مسایل ذکر شده را در یک مدت زمان معقول محاسبه کرد.

تابع چکیده‌ساز MD4، یک تابع چکیده‌ساز اختصاصی[۱۳] است که در سال ۱۹۹۱ توسط رایوست[۱۴] معرفی شد[۱۲]. این تابع اساس طراحی توابع چکیده‌ساز پرکاربردی، مثل MD5، SHA-0، SHA-1، خانواده SHA-2 و خانواده RIPEMD را تشکیل داد که به توابع چکیده‌ساز خانواده MD4 معروف هستند. با توجه به اهمیت امنیت توابع چکیده‌ساز، مقاله‌های متعددی در زمینه تحلیل این توابع منتشر شد. برای مثال در سال ۱۹۹۶، دوبرتین یک حمله برخورد روی تابع چکیده‌ساز MD4 اعمال کرد و توانست با پیچیدگی محاسباتی 2^{22} بار محاسبه این تابع، برای آن یک برخورد پیدا کند[۵]. همچنین در سال ۱۹۹۸، چابود و ژو یک حمله برخورد روی تابع چکیده‌ساز SHA-0 با پیچیدگی محاسباتی 2^{61} بار محاسبه این تابع مطرح کردند[۳]. روش حمله چابود و ژو، اساس حملاتی را تشکیل داد که از آنها می‌توان به نام "تحلیل با استفاده از خطی‌سازی" یاد کرد. این روش توسط رایمن و اوسوالد، برای تحلیل تابع چکیده‌ساز SHA-1 [۱۴] تعمیم و سپس توسط پرامستالر و همکارانش بهبود داده شد[۱۰]. کار مشابه دیگر نیز برای یافتن الگوهای تفاضلی بهینه، با

[1] Confidentiality
[2] Integrity
[3] Availability
[4] Cryptography
[5] Cryptographic Primitives
[6] Digital Signature
[7] Authentication
[8] Hash Family
[9] Psudorandom
[10] Preimage Resistant
[11] 2-Preimagr Resistant
[12] Collision Resistant
[13] Dedicated Hash Function
[14] Rivest

استفاده از خطی‌سازی روی SHA-1، در [۹] صورت گرفته است. همچنین بریر و همکارانش برای تحلیل توابع چکیده‌ساز CubeHash و MD6، از خطی‌سازی استفاده کردند[۲] و سپس کارشان روی تابع چکیده‌ساز CubeHash بهبود داده شد[۶]. از جمله کارهای دیگر که در زمینه تحلیل با استفاده از خطی‌سازی صورت گرفته است، می‌توان به تحلیل تابع چکیده‌ساز SIMD، با استفاده از این روش تحلیل اشاره کرد [۸].

در زمینه تحلیل تابع چکیده‌ساز MD4، می‌توان به کار دوبرتین [۵] و وانگ و همکارانش [۱۵] اشاره کرد که روش‌های تحلیل آنها، متفاوت با روش خطی‌سازی است و تاکنون از روش خطی‌سازی برای تحلیل تابع چکیده‌ساز MD4 استفاده نشده است. هدف اصلی در این مقاله این است که روش تحلیل با استفاده از خطی‌سازی روی تابع چکیده‌ساز MD4 تعمیم داده شود. دلایل انتخاب تابع چکیده‌ساز MD4، علاوه بر دلیل ذکر شده عبارت است از:

۱- تابع چکیده‌ساز MD5 [۱۳]، یک نسخه بهبود یافته از تابع چکیده‌ساز MD4 است که شباهت‌های زیادی با این تابع دارد. بنابراین اگر بتوان با استفاده از خطی سازی، حمله موفقیت‌آمیزی روی تابع چکیده‌ساز MD4 اعمال کرد، می‌توان این حمله را برای تحلیل تابع چکیده‌ساز MD5 نیز تعمیم داد.

۲- با توجه به شباهت زیاد تابع چکیده‌ساز MD4، با توابع چکیده‌ساز MD4 بسط یافته [۱۲] و RIPEMD [۱۱]، می‌توان تحلیل MD4 با استفاده از خطی‌سازی را برای تحلیل این دو تابع و همچنین توابع چکیده‌سازی که بهبودیافته RIPEMD هستند، تعمیم داد.

۳- تفاوت اصلی توابع چکیده‌ساز SHA-0 و SHA-1 (که برای تحلیل آنها از خطی‌سازی استفاده شده است)، با تابع چکیده‌ساز MD4 این است که در عمل‌های مرحله‌ای توابع فشرده‌ساز[1] دو تابع SHA-0 و SHA-1 از ثابت‌های چرخشی با مقادیر یکسان استفاده می‌شود، اما در عمل‌های مرحله‌ای تابع فشرده‌ساز MD4، از ثابت‌های چرخشی با مقادیر متفاوت استفاده می‌شود. علاوه بر این، عمل بسط پیام در این توابع با استفاده از روابط بازگشتی صورت می‌گیرد، در حالی‌که این عمل برای تابع چکیده‌ساز MD4، با استفاده از جایگشت‌های دوری انجام می‌شود (مفهوم تابع فشرده‌ساز و عمل مرحله‌ای تابع فشرده‌ساز و عمل بسط پیام در بخش بعد روشن می‌شود). با تحلیل تابع چکیده‌ساز MD4 با استفاده از خطی‌سازی، می‌توان نقش ثابت‌های چرخشی با مقادیر متفاوت و عمل بسط پیام با استفاده از روابط بازگشتی را در امنیت این تابع مشاهده کرد.

۴- پس از تحلیل موفقیت‌آمیز تابع چکیده‌ساز MD4 با روش مذکور، می‌توان نتایج کار را با استفاده از روش‌های بهبودی، مثل روش بیت-های خنثی بیهام و چن [۱] بهبود داد.

[1] Compression Function

در این مقاله، ابتدا دو نمونه ساده شده از تابع چکیده‌ساز MD4، به اسم‌های توابع چکیده‌ساز LMD4[2] (یک نوع خطی از تابع چکیده‌ساز MD4) و CMD4[3] (یک نوع ساده شده از تابع چکیده‌ساز MD4) تحلیل می‌شود. سپس نتایج حاصل از تحلیل دو تابع مذکور، روی خود تابع چکیده‌ساز MD4 بررسی می‌شود. برای تحلیل توابع چکیده‌ساز LMD4 و CMD4، ماتریس تبدیل LMD4 را به‌دست آورده و با استفاده از این ماتریس و حل یک دستگاه معادلات (با یک روش ابتکاری)، چهار الگوی تفاضلی منجر به برخورد برای LMD4 پیدا می‌شود. چون LMD4 خطی است، با استفاده از هر یک از این الگوها، می‌توان به طور قطعی، برای آن یک برخورد پیدا کرد.

در ادامه با استفاده از یکی از الگوهای تفاضلی ذکرشده، یک حمله برخورد موفقیت‌آمیز روی تابع چکیده‌ساز CMD4، با پیچیدگی 2^{35} بار محاسبه این تابع، اعمال می‌شود (یعنی اگر 2^{35} پیام تصادفی انتخاب شده و زوج متناظر آنها با مقدار تفاضل برابر با مقدار تفاضل الگوی تفاضلی پیدا شده، تولید شود و مقدار چکیده‌سازی هر کدام از این زوج‌ها محاسبه شود، می‌توان انتظار داشت که حداقل مقدار چکیده-سازی یکی از زوج‌ها با هم برابر باشند. به عبارت دیگر زوج پیام مذکور با هم برخورد داشته باشند). در نهایت الگوهای تفاضلی، برای تحلیل تابع چکیده‌ساز MD4 بررسی می‌شوند.

این مقاله از ۶ بخش تشکیل شده است. در بخش ۲، توابع چکیده‌ساز MD4، LMD4 و CMD4 توصیف می‌شوند. در بخش ۳، برای تابع چکیده‌ساز LMD4 چند الگوی تفاضلی منجر به برخورد و با استفاده از یکی از این الگوهای تفاضلی، یک برخورد به دست آورده می‌شود. در بخش ۴، با استفاده از یکی از الگوهای تفاضلی حاصل در بخش ۳، یک حمله برخورد موفقیت‌آمیز روی تابع چکیده‌ساز CMD4 اعمال می‌شود. در بخش۵، استفاده از الگوهای تفاضلی ذکر شده برای تحلیل تابع چکیده‌ساز MD4 بررسی می‌شود. بخش ۶ نیز به نتیجه‌گیری اختصاص یافته است.

۲. توابع چکیده‌ساز MD4، LMD4 و CMD4

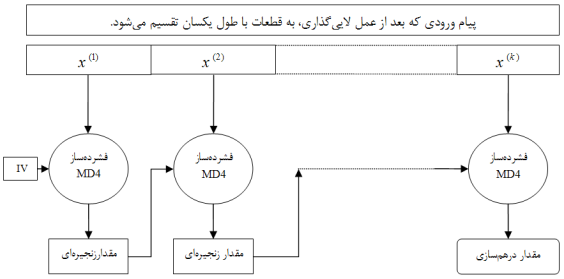
۲-۱. توصیف تابع چکیده‌ساز MD4

تابع چکیده‌ساز MD4، هر پیام با طول دلخواه را به یک مقدار چکیده‌سازی ۱۲۸ بیتی نگاشت می‌کند. برای این کار، از یک عمل فشرده‌سازی در یک ساختار تکرار استفاده شده و تابع فشرده‌ساز MD4 در ساختار تکرار مرکل- دمگارد [۷ و۴] به کار می‌رود (در واقع یک تابع چکیده‌ساز و به طور خاص، تابع چکیده‌ساز MD4، دو جزو اصلی دارد که عبارتند از ساختار تکرار و تابع فشرده‌ساز که به‌طور ساده، نگاشتی مثل $c:\{0,1\}^n\times\{0,1\}^m\to\{0,1\}^n$ تعریف می‌شود).

[2] Linear MD4
[3] Changed MD4

در تابع چکیده‌ساز MD4، ابتدا روی پیام ورودی، عمل لایه‌گذاری[1] صورت می‌گیرد تا طول پیام لایه‌گذاری شده مضربی از ۵۱۲ بیت شود (چون عمل لایه‌گذاری تأثیری در حمله مطرح شده در این مقاله ندارد، از توضیح آن صرف‌نظر می‌شود). سپس پیام لایه‌گذاری شده به قطعه پیام‌های ۵۱۲ بیتی تقسیم می‌شود. هر قطعه پیام ۵۱۲ بیتی با استفاده از تابع فشرده‌ساز MD4 به یک مقدار زنجیره‌ای ۱۲۸ بیتی نگاشت می‌شود (شکل (۱)). در الگوریتم تابع چکیده‌سازMD4، از یک مقدار آغازین ۱۲۸ بیتی استفاده می‌شود که با استفاده از ثابت‌های ۳۲ بیتی و به‌صورت زیر تعریف می‌شود (در این مقاله برای نشان دادن محتویات ثابت‌ها و کلمات از نماد هگزا استفاده شده است):

A=0x67452301, B=0xefcdab89, C=0x98badcfe, D=0x10325476



شکل۱. چکیده‌سازی، با استفاده از تابع فشرده‌ساز MD4 در ساختار تکرار مرکل- دمگارد (تابع چکیده‌ساز MD4)

تابع فشرده‌ساز MD4. تابع فشرده‌ساز MD4، یک قطعه پیام ۵۱۲ بیتی از پیام لایه‌گذاری شده و یک مقدار آغازین ۱۲۸ بیتی را به عنوان ورودی گرفته و یک خروجی ۱۲۸ بیتی تولید می‌کند.

به عبارت دیگر، تابع فشرده‌ساز MD4 یک نگاشت مثل $c:\{0,1\}^{128}\times\{0,1\}^{512}\rightarrow\{0,1\}^{128}$ است. این تابع سه دور دارد که هر دور آن شامل ۱۶ مرحله است.

در هر دور این تابع، از یک تابع بولی متفاوت، مطابق جدول (۱) استفاده می‌شود. برای یک قطعه ورودی ۵۱۲ بیتی مثل $x^{(j)}=\langle X^{(0)},...,X^{(15)}\rangle$ ($X^{(i)}$ ها، $0\le i\le 15$، کلمات ۳۲بیتی هستند و $x^{(j)}$، $j\le 1$، نشان دهنده یک قطعه ۵۱۲ بیتی از پیام لایه‌گذاری شده x است. با توجه به عمل لایه‌گذاری پیام در تابع چکیده‌ساز MD4، حداکثر اندازه پیام ورودی برای این تابع می‌بایست کمتر از 2^{64} باشد که در این‌صورت، کران بالای j، کمتر از 2^{55} خواهد بود. چون این عدد بزرگ است و یا می‌توان با یک قرارداد ساده از محدودیت طول پیام ورودی چشم‌پوشی کرد، بنابراین در عمل حرفی از کران بالای j زده نمی‌شود)، تابع فشرده‌ساز MD4 به‌صورت زیر عمل می‌کند:

۱- فرض کنید A,B,C,D چهار ثبات مورد استفاده در تابع فشرده‌ساز MD4 باشند. اگر $x^{(j)}$ قطعه اول از x (برای فشرده‌سازی) باشد، این ثبات‌ها با مقدار آغازی تابع چکیده‌سازMD4، مقداردهی اولیه می‌شوند. در غیر این‌صورت، آنها با استفاده از خروجی حاصل از فشرده‌سازی قطعه پیام قبلی مقداردهی می‌شوند. بعد از مقداردهی اولیه ثبات‌های ذکر شده، مقادیر آنها به‌ترتیب در $A^{(0)},B^{(0)},C^{(0)},D^{(0)}$ کپی می‌شوند.

۲- برای $0\le i\le 47$، عمل مرحله‌ای زیر، ۴۸ بار تکرار می‌شود تا مقادیر $A^{(i)},B^{(i)},C^{(i)},D^{(i)}$ به‌روز شوند.

$$A^{(i+1)}=D^{(i)}$$
$$B^{(i+1)}=(A^{(i)}+f^{(i)}(B^{(i)},C^{(i)},D^{(i)})+W^{(i)}+K^{(i)})^{<<<s_i}$$
$$C^{(i+1)}=B^{(i)}$$
$$D^{(i+1)}=C^{(i)}$$

در این عمل مرحله‌ای، تابع $f^{(i)}$ و ثابت جمعی $K^{(i)}$ مطابق جدول (۱) و ثابت‌های چرخشی $s^{(i)}$ مطابق جدول (۲) استفاده می‌شوند. $W^{(i)}$ ها نیز کلمات ۳۲ بیتی هستند که با استفاده از عمل بسط پیام روی ۱۶ کلمه قطعه پیام $x^{(j)}$ حاصل می‌شوند. عمل بسط پیام در تابع فشرده‌ساز MD4، توسط جایگشت‌های دوری و مطابق جدول (۳) صورت می‌گیرد که در آن k، $0\le k\le 2$، شماره دور در تابع فشرده‌ساز MD4 است(نماد +، جمع در پیمانه 2^{32} و نماد $<<<s_i$، چرخش بیتی به مقدار s_i، به سمت چپ است).

۳- مقادیر به‌روز شده $A^{(i)},B^{(i)},C^{(i)},D^{(i)}$ بعد از مرحله ۴۷ام به ترتیب با مقادیر زنجیره‌ای ورودی تابع فشرده‌ساز، جمع پیمانه‌ای می‌شوند تا مقادیر زنجیره نهایی برای قطعه ورودی در نظر گرفته شده، به‌صورت زیر تولید شوند:

$$A=A^{(48)}+A^{(0)},B=B^{(48)}+B^{(0)},$$
$$C=C^{(48)}+C^{(0)},D=D^{(48)}+D^{(0)}$$

اگر $x^{(j)}$ قطعه آخر از پیام x باشد، مقدار چکیده‌سازی x برابر A||B||C||D خواهد بود (که در آن منظور از نماد ||، الحاق دو مقدار به هم است). در غیر این‌صورت، فرآیند بالا برای قطعه ۵۱۲ بیتی بعدی از x، با مقادیر زنجیره‌ای A,B,C,D محاسبه شده جدید، تکرار می‌شود.

جدول ۱. توابع بولی و ثابت‌های جمعی برای تابع فشرده‌ساز MD4

مرحله i	تابع $f^{(i)}$	ثابت $K^{(i)}$
0,…,15	IF: $(X\wedge Y)\vee(\bar{X}\wedge Z)$	0x00000000
16,…,31	MAJ: $(X\wedge Y)\vee(X\wedge Z)\vee(Y\wedge Z)$	0x5a827999
32,…,47	XOR: $X\oplus Y\oplus Z$	0x6ed9eba1

[1] Padding

جدول ۲. ثابت‌های چرخشی در تابع فشرده‌ساز MD4

j	$s^{(4j)}$	$s^{(4j+1)}$	$s^{(4j+2)}$	$s^{(4j+3)}$
0,…,3	3	7	11	19
4,…,7	3	5	9	13
8,…,11	3	9	11	15

۲-۲. توصیف تابع چکیده‌ساز LMD4

تابع چکیده‌ساز LMD4، یک تقریب خطی از تابع چکیده‌ساز MD4 است. در تابع فشرده‌ساز MD4، جمع پیمانه‌ای و توابع بولی در عمل‌های مرحله‌ای تابع فشرده‌ساز MD4، دو عامل ایجاد خاصیت غیرخطی در تابع چکیده‌ساز MD4 هستند. تابع چکیده‌ساز LMD4، همان تابع چکیده‌ساز MD4 در نظر گرفته می‌شود، با این تفاوت که در تابع فشرده‌ساز آن به جای جمع پیمانه‌ای از عمل‌گر XOR و به جای توابع بولی IF و MAJ از تابع بولی XOR استفاده می‌شود.

۲-۳. توصیف تابع چکیده‌سازCMD4

تابع چکیده‌سازCMD4 همان تابع چکیده‌سازMD4 است، با این تفاوت که در دور اول تابع فشرده‌ساز آن به جای تابع بولی IF از تابع بولی XOR استفاده می‌شود.

تذکر۱: حمله مطرح شده در این مقاله، یک حمله برخورد به توابع چکیده‌ساز LMD4، CMD4 و MD4، با یک بار تکرار تابع فشرده‌ساز آنها در ساختار تکرار مرکل- دمگارد است. به همین خاطر در ادامه، توابع چکیده‌ساز ذکر شده تنها با یک تکرار تابع فشرده‌ساز در ساختار تکرار مرکل- دمگارد در نظر گرفته می‌شوند.

۳. به‌دست آوردن برخورد برای تابع چکیده‌ساز LMD4

ابتدا تابع فشرده‌ساز LMD4 را در دو بخش بسط پیام و تابع به‌روز رسانی حالت (شامل عمل‌های مرحله‌ای) در نظر بگیرید. عمل بسط پیام در این تابع، یک قطعه ۵۱۲ بیتی، یا ۱۶ کلمه ۳۲ بیتی را به عنوان ورودی گرفته و یک قطعه ۱۵۳۶ بیتی، یا ۴۸ کلمه ۳۲ بیتی را به‌عنوان خروجی تولید می‌کند. چون این عمل خطی است، بنابراین می‌توان برای آن یک ماتریس تبدیل مثل $E_{512\times1536}$ به‌دست آورد. با فرض اینکه $X_{1\times512}$ نمایش ماتریسی قطعه پیام ۵۱۲ بیتی باشد، رابطه (۱) را می‌توان نوشت:

$$W_{1\times1536} = X_{1\times512}.E_{512\times1536} \qquad (۱)$$

نمایش ماتریسی قطعه ۱۵۳۶ بیتی حاصل از بسط پیام خواهد بود (که در آن منظور از "." ضرب ماتریسی می‌باشد). حال تابع به‌روز رسانی حالت در تابع فشرده‌ساز LMD4 را در نظر بگیرید. چون این تابع نیز خطی است، می‌توان آن را با یک ماتریس مثل $S_{1536\times128}$، یک ماتریس مثل $T_{128\times128}$ و یک بردار مثل $K_{1\times128}$ در نظر گرفت که در آن $S_{1536\times128}$ ماتریس تبدیل حالت تابع فشرده‌ساز، $T_{128\times128}$ ماتریس تبدیل مقدار آغازین تابع فشرده‌ساز و $K_{1\times128}$ نیز ماتریس تبدیل ثابت‌های جمعی به‌کار رفته در تابع فشرده‌ساز است. با استفاده از ماتریس‌های S، T و K، برای یک قطعه حاصل از بسط پیام، می‌توان یک بردار خروجی مثل $O_{1\times128}$ را به‌صورت زیر تولید کرد:

$$O_{1\times128} = W_{1\times1536}\cdot S_{1536\times128} \oplus IV_{1\times128}\cdot T_{128\times128} \oplus K_{1\times128}$$

که در آن $IV_{1\times128}$ نمایش برداری مقدار آغازین تابع چکیده‌ساز LMD4 است و تبدیل خطی اعمال شده روی آن، با ماتریس T در نظر گرفته شده است. بردار K نیز تبدیل خطی اعمال شده روی ثابت‌های جمعی LMD4 است. با توجه به رابطه (۱) می‌توان نوشت:

$$O_{1\times128} = X_{1\times512}\cdot E_{512\times1536}\cdot S_{1536\times128} \oplus IV_{1\times128}\cdot T_{128\times128} \oplus K_{1\times128} \qquad (۲)$$

در واقع، برای یک پیام یک قطعه‌ای، مثل X، $O_{1\times128}$ نمایش برداری مقادیر ثبات‌های تابع فشرده‌ساز LMD4، در انتهای مرحله ۴۷ آن می‌باشد.

حال فرض کنید X_1 و $X_2 = X_1 \oplus \mu$، دو پیام یک قطعه‌ای با تفاضل μ باشند که با استفاده از تابع چکیده‌سازLMD4 با هم برخورد دارند. یعنی با توجه به رابطه (۲) داریم:

$$X_1\cdot E\cdot S \oplus IV\cdot T \oplus K = (X_1 \oplus \mu)\cdot E\cdot S \oplus IV\cdot T \oplus K$$

این رابطه را می‌توان به شکل زیر ساده کرد:

$$X_1\cdot E\cdot S \oplus IV\cdot T = (X_1 \oplus \mu)\cdot E\cdot S \oplus IV\cdot T \Rightarrow$$
$$X_1\cdot E\cdot S = (X_1 \oplus \mu)\cdot E\cdot S \Rightarrow$$
$$X_1\cdot E\cdot S = X_1\cdot E\cdot S \oplus \mu\cdot E\cdot S$$

از رابطه اخیر نیز می‌توان نتیجه گرفت که دو پیام X_1 و $X_2 = X_1 \oplus \mu$، با هم برخورد دارند اگر و تنها اگر $\mu\cdot E\cdot S = 0$.

جدول ۳. جایگشت‌های دوری برای عمل بسط پیام در MD4

$\sigma_k(i)$		i															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
k	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	1	0	4	8	12	1	5	9	13	2	6	10	14	3	7	11	15
	2	0	8	4	12	2	10	6	14	1	9	5	13	3	11	7	15

به‌عبارت دیگر، با در نظر گرفتن $H_{512\times128} = E_{512\times1536}.S_{1536\times128}$، می‌توان گفت دو پیام X_1 و $X_2 = X_1 \oplus \mu$ با هم برخورد دارند، اگر و تنها اگر

$$\mu_{1\times512} \cdot H_{512\times1536} = 0 \qquad (۳)$$

رابطه (۳)، نشان‌دهنده یک دستگاه معادلات است که هر پاسخ از آن (یعنی یک جواب برای μ)، یک الگوی تفاضلی ورودی برای تابع چکیده‌ساز LMD4 خواهد بود. یعنی می‌توان یک پیام تصادفی مانند X_1 پیدا کرد و سپس با استفاده از پاسخ μ، $X_2 = X_1 \oplus \mu$ را محاسبه کرد، به طوری که با توجه به توضیحات بالا، در استفاده از تابع چکیده‌ساز LMD4، مقدار چکیده‌سازی X_1 و X_2 با هم یکسان خواهند بود. به‌عبارت دیگر، دو پیام X_1 و X_2، برای تابع چکیده‌ساز LMD4، یک برخورد نتیجه خواهند داد.

بعد از بسط الگوی تفاضلی ورودی حاصل از پاسخ μ با استفاده از عمل بسط پیام در LMD4، می‌توان یک الگوی تفاضلی منجر به برخورد برای LMD4 به دست آورد که در این الگو، تفاضلات در هر مرحله از تابع فشرده‌ساز مشخص می‌باشد. اما پاسخی از دستگاه معادلات مذکور می‌تواند برای تحلیل تابع چکیده‌ساز CMD4، نتیجه بهتری داشته باشد که منجر به یک الگوی تفاضلی با مینیمم وزن همینگ[1] شود(منظور از وزن همینگ الگوی تفاضلی، تعداد یک‌ها در این الگو می‌باشد)[۱۴].

از طرف دیگر چون عمل بسط پیام در تابع فشرده‌ساز LMD4، توسط جایگشت‌های دوری صورت می‌گیرد، بنابراین پاسخی از دستگاه معادلات (۳) می‌تواند منجر به یک الگوی تفاضلی با وزن همینگ کم یا مینیمم شود که خود پاسخ، حداقل وزن همینگ را داشته باشد. در این مقاله، برای پیدا کردن الگوی تفاضلی با وزن همینگ کم از یک روش جستجوی تصادفی، همراه با یک روش ابتکاری، به‌صورت زیر استفاده شده است:

۱- فرض کنید $\mu_{1\times512} = [x_0, x_1, ..., x_{511}]$ باشد. ابتدا متغیرهای $\mu_{1\times512}$ یک در میان برابر صفر قرار داده می‌شود. برای مثال فرض کنید برای iهای زوج در $\mu_{1\times512}$، $x_i = 0$ در نظر گرفته شود. مجموعه x_iها با i فرد را با $\alpha_{1\times256} = [x_1, x_3, x_5, ..., x_{511}]$ در نظر بگیرید.

۲- به طور مشابه، متغیرهای $\alpha_{1\times256}$، یک در میان برابر صفر قرار داده می‌شود. برای مثال، فرض کنید در $\alpha_{1\times256}$، برای $0 \le i \le 127$، $x_{(4i+1)} = 0$ در نظر گرفته شود.

متغیرهای باقیمانده را با $\beta_{1\times128} = [x_3, x_7, x_{11}, ..., x_{511}]$ در نظر بگیرید.

۳- سعی می‌شود، یک مقدار با حداقل وزن همینگ برای $\beta_{1\times128}$ به‌دست آورده شود. به‌طوری که با در نظر گرفتن این پاسخ همراه با متغیرهایی از $\mu_{1\times512}$ که در مراحل ۱و ۲ برابر صفر قرار داده شده‌اند، بتوان یک پاسخ برای دستگاه معادلات (۳) به‌دست آورد.

روشن است که با استفاده از روش بالا، وزن همینگ پاسخ $\mu_{1\times512}$ پیدا شده با وزن همینگ مقدار حاصل برای $\beta_{1\times128}$ یکسان خواهد بود. در این مقاله، با استفاده از روش جستجوی ذکر شده، بهترین پاسخ‌هایی که برای دستگاه معادلات(۳) پیدا شد، چهار پاسخ با وزن همینگ ۴۰ می‌باشد که در جدول (۴) آورده شده است.

در این جدول، هر پاسخ μ_j با ۱۶ کلمه ۳۲ بیتی $\mu^{(i)}$، به‌صورت $\mu_j = \mu^{(0)} \| ... \| \mu^{(15)}$ در نظر گرفته شده و $\mu^{(i)}$ها نیز با نماد هگزا نشان داده شده‌اند. با استفاده از هر یک از این پاسخ‌ها، یک الگوی تفاضلی با وزن ۱۲۰، برای LMD4 تولید می‌شود (چون اگر یکی از این پاسخ‌ها در عمل بسط پیام LMD4، که با جایگشت‌های دوری انجام می‌شود، وارد شود، وزن همینگ آن سه برابر می‌شود).

چون در تابع چکیده‌ساز LMD4، همه چیز خطی است، بنابراین با استفاده از یکی از پاسخ‌های جدول (۴)، می‌توان به طور قطعی یک برخورد به‌دست آورد.

جدول ۴. چهار پاسخ با وزن همینگ ۴۰ برای دستگاه معادلات (۳)

$\mu^{(i)}$	پاسخ μ_1	پاسخ μ_2	پاسخ μ_3	پاسخ μ_4
$\mu^{(0)}$	00000000	00000000	00000000	00000000
$\mu^{(1)}$	88888888	44444444	22222222	11111111
$\mu^{(2)}$	88888888	44444444	22222222	11111111
$\mu^{(3)}$	00000000	00000000	00000000	00000000
$\mu^{(4)}$	00000000	00000000	00000000	00000000
$\mu^{(5)}$	00000000	00000000	00000000	00000000
$\mu^{(6)}$	88888888	44444444	22222222	11111111
$\mu^{(7)}$	00000000	00000000	00000000	00000000
$\mu^{(8)}$	00000000	00000000	00000000	00000000
$\mu^{(9)}$	00000000	00000000	00000000	00000000
$\mu^{(10)}$	88888888	44444444	22222222	11111111
$\mu^{(11)}$	00000000	00000000	00000000	00000000
$\mu^{(12)}$	00000000	00000000	00000000	00000000
$\mu^{(13)}$	00000000	00000000	00000000	00000000
$\mu^{(14)}$	00000000	00000000	00000000	00000000
$\mu^{(15)}$	88888888	44444444	22222222	11111111

[1] Hamming Weight

یکی از این برخوردها در مثال ۱ آورده شده است.

مثال ۱: با استفاده از پاسخ μ_1 در جدول (۴)، پیام‌های ۵۱۲ بیتی

21438c50	a8e3a1aa	811915a4	3c61d57e
3554dec7	1b4839ba	ac9cacad	096c2ab0
07957893	0cd7fc43	a0256f49	20384f00
0060be34	3119954e	00550182	a409be7d

و

21438c50	206b2922	09919d2c	3c61d57e
3554dec7	1b4839ba	24142425	096c2ab0
07957893	0cd7fc43	28ade7c1	20384f00
0060be34	3119954e	00550182	2c8136f5

بعد از مرحله ۴۷ام تابع فشرده‌ساز LMD4، مقدار خروجی ۱۲۸ بیتی یکسان زیر را دارند:

e5fd4ee9 dc2ff2be 8c1cb5c7 e82180c3

که در آنها ترتیب کلمات از چپ به راست و بالا به پایین می‌باشد. در نتیجه دو پیام بالا، با توجه به توصیف تابع چکیده‌ساز LMD4، با هم برخورد خواهند داشت (چون در توصیف تابع فشرده‌ساز LMD4، خروجی تابع فشرده‌ساز از جمع پیمانه‌ای مقدار خروجی در انتهای مرحله ۴۷ تابع فشرده‌ساز با مقدار آغازین آن حاصل می‌شود و با توجه به اینکه مقدار آغازین مورد استفاده برای هر دو پیام یکسان در نظر گرفته شده است، بنابراین می‌توان گفت که در استفاده از تابع چکیده‌ساز LMD4، برای چکیده‌سازی پیام‌ها، این دو پیام با هم‌دیگر برخورد خواهند داشت).

در ادامه با استفاده از یکی از پاسخ‌های دستگاه معادلات(۳) در جدول (۴)، یک حمله برخورد روی تابع چکیده‌سازCMD4 اعمال می‌شود. برای این منظور پاسخ μ_1 در جدول (۴) را در نظر بگیرید. حال باید احتمال موفقیت پاسخ μ_1 برای به دست آوردن برخورد در تابع چکیده‌سازCMD4 بررسی شود.

چون در CMD4، همه چیز خطی نیست، بنابراین باید بررسی شود که (در استفاده از μ_1) با چه احتمالی تابع چکیده‌ساز CMD4، مثل تابع چکیده‌ساز LMD4 عمل می‌کند. یا به طور معادل، با چه احتمالی می‌توان دو پیام X_1 و $X_2 = X_1 \oplus \mu_1$ را طوری پیدا کرد که با استفاده از تابع چکیده‌ساز CMD4 با هم برخورد داشته باشند.

برای این کار، می‌بایست در هر مرحله از تابع فشرده‌ساز CMD4، احتمال اینکه عمل مرحله‌ای در این تابع فشرده‌ساز، مثل عمل مرحله‌ای تابع فشرده‌ساز LMD4 عمل کند، محاسبه شود. بنابراین لازم است تا μ_1 با استفاده از عمل بسط پیام در CMD4 بسط داده شود تا تفاضلات برای هر مرحله مشخص باشد. قطعه حاصل از بسط μ_1 با استفاده از عمل بسط پیام CMD4، الگوی تفاضلی λ نامیده می‌شود. الگوی تفاضلی λ یک قطعه شامل ۴۸ کلمه ۳۲ بیتی است که هر کلمه i آن، $0 \le i \le 47$، معادل با تفاضل ورودی در مرحله i تابع فشرده‌ساز CMD4 است.

۴. تحلیل تابع چکیده‌ساز CMD4

برای این‌که بتوان با استفاده از الگوی تفاضلی λ، برای تابع چکیده‌ساز CMD4 یک برخورد به‌دست آورد، لازم است در استفاده از این الگوی تفاضلی، تابع چکیده‌ساز CMD4 همانند تقریب خطی خود، یعنی تابع چکیده‌ساز LMD4 عمل کند. به عبارت دیگر لازم است تا تابع بولی MAJ یک رفتار تفاضلی یکسان با رفتار تفاضلی تابع بولی XOR داشته و جمع پیمانه‌ای نیز مثل عمل‌گر XOR عمل کند.

برای این امر نیز می‌بایست برخی شرایط روی مقادیر پیام‌ها و مقادیر ثبات‌ها برآورده شود. این شرایط با توجه به الگوی تفاضلی و تفاضلات بیتی در متغیرهای ورودی تابع بولی MAJ و تابع بولی XOR (تفاضلات ورودی و خروجی) و عدم تولید رقم نقلی[1] تعیین می‌شوند. با توجه به الگوی تفاضلی λ، تفاضلات ورودی و خروجی در هر مرحله از تابع فشرده‌سازLMD4، در جدول(۵) آورده شده است. برای مثال، با توجه به جدول(۵)، تفاضلات مربوط به بیت ۲۹ مقادیر پیام‌ها و ثبات‌ها، در مرحله ۳۰ تابع فشرده‌ساز LMD4 برابر با:

$\Delta^{\oplus}W_{29}^{30} = 0$، $\Delta^{\oplus}A_{29}^{30} = 0$، $\Delta^{\oplus}B_{29}^{30} = 0$، $\Delta^{\oplus}C_{29}^{30} = 1$

و $\Delta^{\oplus}D_{29}^{30} = 1$ است(منظور از $\Delta^{\oplus}Z_j^{(i)} = Z_j^{(i)} \oplus Z_j'^{(i)}$، تفاضل XOR در بیت j $Z^{(i)}$ و $Z'^{(i)}$ است).

چون در مرحله ۳۰ از تابع فشرده‌ساز CMD4، تابع بولی MAJ استفاده شده است، بنابراین ابتدا باید مطمئن شد که تابع بولی MAJ مثل تابع بولی XOR رفتار کند. با توجه به جدول(۶) شرط لازم برای این امر این است که $C_{29}^{(30)} \oplus D_{29}^{(30)} = 1$ باشد. از طرف دیگر، با توجه به عمل مرحله‌ای در تابع فشرده‌ساز CMD4، $C_{29}^{(30)} = B_{29}^{(29)}$ و $D_{29}^{(30)} = B_{29}^{(28)}$ است. بنابراین شرط لازم برای این‌که برای بیت ۲۹ در مرحله ۳۰ تابع فشرده‌ساز CMD4، تابع MAJ، مثل تابع XOR عمل کند، شرط $B_{29}^{(29)} \oplus B_{29}^{(28)} = 1$ خواهد بود.

چون برای محاسبه $B^{(29)}$ و $B^{(28)}$ از کلمات $W^{(28)}$ و $W^{(27)}$ استفاده شده است، بنابراین می‌توان این کلمات را طوری پیدا کرد که شرط مذکور برآورده شود. مشابه با روش بالا، تمام شرایط لازم برای این‌که تابع MAJ در دور دوم تابع فشرده‌ساز CMD4، مثل تابع XOR رفتار کند و هیچ رقم نقلی نیز تولید نشود، مشخص می‌شود. توجه شود چون دور اول تابع فشرده‌ساز CMD4 و LMD4 یکسان هستند، بنابراین برای دور اول تابع فشرده‌ساز CMD4، هیچ شرطی لازم نیست. از طرف دیگر عمل بسط پیام در تابع فشرده‌ساز CMD4، توسط جایگشت‌های دوری انجام می‌شود، بنابراین مشابه با روش چابود و ژو [۳] یا روش اصلاح پیام وانگ[۱۵]، می‌توان ابتدا کلماتی از یک قطعه‌پیام ورودی را طوری پیدا کرد که تمام شرایط لازم در دور دوم برآورده شوند و سپس کار جستجوی برخورد را شروع کرد.

[1] Carry

بنابراین، در محاسبه پیچیدگی حمله روی تابع چکیده‌ساز CMD4، می‌توان تعداد شرایط برای دور دوم تابع فشرده‌ساز CMD4 را نادیده گرفت. اما تعداد شرایط برای دور سوم تابع فشرده‌ساز CMD4 مهم است. چون در دور سوم تابع فشرده‌ساز CMD4، تابع بولی XOR استفاده می‌شود، بنابراین تنها باید شرایط لازم برای این‌که جمع پیمانه‌ای مثل عمل‌گر XOR عمل کند، مشخص شود. توجه شود برای حالتی که $\Delta^{\oplus}W_j^{(i)}=0$ است، جمع پیمانه‌ای مثل XOR عمل خواهد کرد و لازم نیست تا شرطی برآورده شود. اما برای محل بیت‌هایی که در آن‌ها تفاضل غیر صفر وجود دارد، یعنی برای حالتی که $\Delta^{\oplus}W_j^{(i)}\neq 0$ است، لازم است تا شرایطی تعیین شود.

برای مشخص کردن این شرایط، ابتدا فرض کنید که خروجی تابع $f^{(i)}$، برای سه متغیر ورودی $B^{(i)}$، $C^{(i)}$ و $D^{(i)}$ با $f^{(i)*}$ تعیین شده باشد. در جدول (۵) مرحله ۳۶ از تابع فشرده‌ساز CMD4 در نظر گرفته می‌شود، برای بیت سوم در این مرحله، تفاضلات زیر وجود دارد:

$$\Delta^{\oplus}W_3^{(36)}=1 \quad ,\Delta^{\oplus}A_3^{(36)}=0 \quad ,\Delta^{\oplus}B_3^{(36)}=0 \quad ,$$
$$\Delta^{\oplus}C_3^{(36)}=0 \quad ,\Delta^{\oplus}D_3^{(36)}=1$$

چون تابع $f^{(36)}$، تابع XOR است، بنابراین:

$$\Delta^{\oplus}f_3^{*(36)}=f_3^{*(36)}\oplus f_3'^{*(36)}=\Delta^{\oplus}B_3^{(36)}\oplus\Delta^{\oplus}C_3^{(36)}\oplus\Delta^{\oplus}D_3^{(36)}=1$$

خواهد بود. با توجه به $\Delta^{\oplus}W_3^{(36)}=1$ و $\Delta^{\oplus}f_3^{*(36)}=1$، حالت‌های ممکن زیر می‌تواند برای W_3^{36}، $W_3'^{36}$، $f_3^{*(36)}$ و $f_3'^{*(36)}$ روی دهد.

$f_3^{*(36)}$	$f_3'^{*(36)}$
0	1
1	0

W_3^{36}	$W_3'^{36}$
0	1
1	0

با توجه به این حالت‌ها برای این‌که در بیت سوم از مرحله سی و ششم، رقم نقلی تولید نشود، لازم است تا $f_3^{*(36)}\oplus W_3^{36}=1$ باشد که در آن $f_3^{*(36)}=B_3^{36}\oplus C_3^{36}\oplus D_3^{36}$ است. بنابراین شرط لازم برای این‌که در محل مورد نظر رقم نقلی تولید نشود $B_3^{36}\oplus C_3^{36}\oplus D_3^{36}\oplus W_3^{36}=1$ خواهد بود که با توجه به عمل مرحله‌ای در تابع فشرده‌ساز CMD4، شرط اخیر را نیز می‌توان به‌صورت زیر نوشت:

$$B_3^{36}\oplus B_3^{35}\oplus B_3^{34}\oplus W_3^{36}=1$$

چون در محاسبه B_3^{34}، B_3^{35} و B_3^{36} از کلمات پیام استفاده می‌شود، بنابراین می‌توان انتظار داشت پیامی یافت شود که این شرط برآورده شود. به طور مشابه برای تمام مراحل دور سوم تابع فشرده‌ساز CMD4، که تفاضل غیر صفر وجود دارد، شرایط تعیین می‌شوند. توجه شود برای بیت‌های ۳۱ در کلمات λ که در دور سوم تابع فشرده‌ساز CMD4 استفاده می‌شوند، شرطی لازم نیست. چون در بیت‌های ۳۱ امکان تولید رقم نقلی وجود ندارد.

جدول ۵. تفاضلات برای مقادیر ثابت‌های دو قطعه حاصل از بسط پیام با تفاضل λ، در تابع چکیده‌ساز LMD4

مرحله i	$\Delta^{\oplus}W^{(i)}$	$\Delta^{\oplus}A^{(i)}$	$\Delta^{\oplus}B^{(i)}$	$\Delta^{\oplus}C^{(i)}$	$\Delta^{\oplus}D^{(i)}$
0	00000000	00000000	00000000	00000000	00000000
1	88888888	00000000	00000000	00000000	00000000
2	88888888	00000000	44444444	00000000	00000000
3	00000000	00000000	66666666	44444444	00000000
4	00000000	00000000	11111111	66666666	44444444
5	00000000	44444444	99999999	11111111	66666666
6	88888888	66666666	55555555	99999999	11111111
7	00000000	11111111	99999999	55555555	99999999
8	00000000	99999999	22222222	99999999	55555555
9	00000000	55555555	bbbbbbbb	22222222	99999999
10	88888888	99999999	aaaaaaaa	bbbbbbbb	22222222
11	00000000	22222222	11111111	aaaaaaaa	bbbbbbbb
12	00000000	bbbbbbbb	11111111	11111111	aaaaaaaa
13	00000000	aaaaaaaa	88888888	11111111	11111111
14	00000000	11111111	11111111	88888888	11111111
15	88888888	11111111	cccccccc	11111111	88888888
16	00000000	88888888	66666666	cccccccc	11111111
17	00000000	11111111	99999999	66666666	cccccccc
18	00000000	cccccccc	44444444	99999999	66666666
19	00000000	66666666	eeeeeeee	44444444	99999999
20	88888888	99999999	aaaaaaaa	eeeeeeee	44444444
21	00000000	44444444	88888888	aaaaaaaa	eeeeeeee
22	00000000	eeeeeeee	11111111	88888888	aaaaaaaa
23	00000000	aaaaaaaa	bbbbbbbb	11111111	88888888
24	88888888	88888888	11111111	bbbbbbbb	11111111
25	88888888	11111111	dddddddd	11111111	bbbbbbbb
26	88888888	bbbbbbbb	dddddddd	dddddddd	11111111
27	00000000	11111111	44444444	dddddddd	dddddddd
28	00000000	dddddddd	aaaaaaaa	44444444	dddddddd
29	00000000	dddddddd	77777777	aaaaaaaa	44444444
30	00000000	44444444	88888888	77777777	aaaaaaaa
31	88888888	aaaaaaaa	22222222	88888888	77777777
32	00000000	77777777	ffffffff	22222222	88888888
33	00000000	88888888	11111111	ffffffff	22222222
34	00000000	22222222	88888888	11111111	ffffffff
35	00000000	ffffffff	22222222	88888888	11111111
36	88888888	11111111	22222222	22222222	88888888
37	88888888	88888888	88888888	22222222	22222222
38	88888888	22222222	11111111	88888888	22222222
39	00000000	22222222	88888888	11111111	88888888
40	88888888	88888888	99999999	88888888	11111111
41	00000000	11111111	00000000	99999999	88888888
42	00000000	88888888	00000000	00000000	99999999
43	00000000	99999999	88888888	00000000	00000000
44	00000000	00000000	88888888	88888888	00000000
45	00000000	00000000	00000000	88888888	88888888
46	00000000	88888888	00000000	00000000	88888888
47	88888888	88888888	00000000	00000000	00000000

تعداد شرایط لازم برای این‌که در تمام مراحل دور سوم تابع فشرده-ساز CMD4، رقم نقلی تولید نشود، برابر ۳۵ شرط خواهد بود. برخلاف شرایط مربوط به دور دوم تابع فشرده‌ساز CMD4، این شرایط به‌صورت احتمالی برآورده می‌شوند. چون شرایط برای بیت‌ها در نظر گرفته شده است، بنابراین احتمال برآورده شدن هر شرط برابر با $\frac{1}{2}$ می‌باشد. در نتیجه احتمال موفقیت الگوی تفاضلی λ، برای به دست آوردن برخورد در تابع چکیده‌ساز CMD4، برابر با احتمال برآورده شدن این شرایط، یعنی 2^{-35} خواهد بود. به‌عبارت دیگر، می‌توان گفت که پیچیدگی حمله برخورد اعمال‌شده روی تابع چکیده‌ساز CMD4، برابر با 2^{35} بار محاسبه تابع چکیده‌ساز CMD4 خواهد بود.

تذکر۲: اگر در حمله برخورد روی تابع چکیده‌ساز CMD4، با استفاده از الگوی تفاضلی λ، از روش اصلاح پیام چند مرحله‌ای وانگ[۱۵] استفاده شود، می‌توان تعدادی از شرایط لازم برای دور سوم را نیز به طور قطعی برآورده و در نتیجه پیچیدگی حمله را کاهش داد.

همچنین، اگر از روش بیت‌های خنثی بیهام و چن[۱] استفاده شود، باز هم می‌توان پیچیدگی حمله را کاهش داد. چون با استفاده از روش بیت‌های خنثی بیهام و چن، می‌توان شرایط را برای بیش از ۱۶ مرحله(تعداد مراحل یک دور از تابع فشرده‌ساز CMD4)، به‌طور قطعی برآورده کرد.

جدول ۶. شرایطی که لازم است تا توابع IF و MAJ رفتار تفاضلی یکسان با رفتار تفاضلی تابع XOR داشته باشند[۱۰].

شرایط لازم برای اینکه توابع IF و MAJ مثل XOR رفتار کنند		تفاضلات ورودی، مبتنی بر بیت‌ها		
تابع MAJ	تابع IF	$\Delta^{\oplus}D_j^{(i)}$	$\Delta^{\oplus}C_j^{(i)}$	$\Delta^{\oplus}B_j^{(i)}$
همیشه	همیشه	0	0	0
$B_j^{(i)} \oplus C_j^{(i)} = 1$	$B_j^{(i)} = 0$	1	0	0
$B_j^{(i)} \oplus D_j^{(i)} = 1$	$B_j^{(i)} = 1$	0	1	0
$C_j^{(i)} \oplus D_j^{(i)} = 1$	هرگز	1	1	0
$C_j^{(i)} \oplus D_j^{(i)} = 1$	$C_j^{(i)} \oplus D_j^{(i)} = 1$	0	0	1
$B_j^{(i)} \oplus D_j^{(i)} = 1$	$B_j^{(i)} \oplus C_j^{(i)} \oplus D_j^{(i)} = 0$	1	0	1
$B_j^{(i)} \oplus C_j^{(i)} = 1$	$B_j^{(i)} \oplus C_j^{(i)} \oplus D_j^{(i)} = 0$	0	1	1
همیشه	$C_j^{(i)} \oplus D_j^{(i)} = 0$	1	1	1

۵. احتمال موفقیت الگوی تفاضلی λ، برای به دست آوردن برخورد در تابع چکیده‌ساز MD4

مشابه با کار تحلیل CMD4، برای این‌که بتوان با استفاده از الگوی تفاضلی λ، برای تابع چکیده‌سازMD4 یک برخورد به‌دست آورد، لازم است تا توابع بولی IF و MAJ، یک رفتار تفاضلی یکسان با رفتار تفاضلی تابع بولی XOR داشته باشند و علاوه بر این، اعمال تفاضل-های موجود در λ روی یک قطعه حاصل از بسط پیام، باعث تولید رقم نقلی نشود، یعنی جمع پیمانه‌ای نیز مثل عمل‌گر XOR عمل کند. برای برقراری شرایط فوق، لازم است تا محدودیت‌هایی روی مقادیر ثابت‌ها و مقادیر پیام‌ها اعمال شوند.

با برآورده شدن این شرایط اطمینان حاصل می‌شود که تابع چکیده-سازMD4 مانند تابع چکیده‌ساز LMD4 عمل می‌کند. برای مشخص کردن این شرایط لازم است تا تفاضلات خروجی در هر مرحله از تابع فشرده‌ساز LMD4 نیز مشخص شوند که این تفاضلات در جدول (۵) آورده شده‌اند.

حال برای مشخص شدن شرایط لازم برای این‌که توابع بولی غیرخطی، مثل تابع XOR عمل کنند، لازم است جدول (۶) مورد توجه قرار گیرد. با توجه به این جدول، برای تابع بولی IF نباید حالتی پیش بیاید که در آن دو ورودی $C_j^{(i)}$ و $D_j^{(i)}$ هم‌زمان تغییر کنند و $B_j^{(i)}$ بدون تغییر باقی بماند. چون در این حالت، تابع بولی IF هرگز نمی‌تواند یک رفتار تفاضلی یکسان با تابع بولی XOR داشته باشد.

در کاربرد الگوهای تفاضلی حاصل، در این مقاله برای به دست آوردن برخورد در تابع چکیده‌سازMD4، حالت ذکر شده برای تابع بولی IF روی می‌دهد. در نتیجه نمی‌توان از این الگوهای تفاضلی موجود برای به دست آوردن برخورد در تابع چکیده‌ساز MD4 استفاده کرد (یعنی احتمال موفقیت هر یک از الگوهای تفاضلی حاصل در این مقاله برای پیدا کردن برخورد در تابع چکیده‌سازMD4 برابر ۰ است) و می‌بایست به دنبال یک الگوی تفاضلی مناسب بود که در استفاده از این الگو، حالت مذکور برای تابع IF پیش نیاید.

۶. نتیجه‌گیری

اصول طراحی تابع چکیده‌ساز MD4، اساس طراحی توابع چکیده‌ساز پرکاربرد امروزی، مثل توابع چکیده‌ساز MD5، SHA-0 و SHA-1 را تشکیل می‌دهد. دو روش مهم برای تحلیل تابع چکیده‌ساز MD4، روش حمله دوبرتین و روش حمله وانگ است، ولی تاکنون برای تحلیل این تابع، از خطی‌سازی استفاده نشده است.

در این مقاله با استفاده از خطی‌سازی، دو نمونه ساده شده از تابع چکیده‌ساز MD4 تحلیل شدند. برای این منظور ابتدا نمونه کاملاً خطی از تابع چکیده‌سازMD4، به اسم تابع چکیده‌سازLMD4 در نظر

گرفته شده و با استفاده از جبرخطی و یک روش جستجوی ابتکاری، چهار الگوی تفاضلی ورودی با وزن همینگ ۴۰ برای این تابع به دست آورده شد. چون در LMD4 همه چیز خطی بود، بنابراین با استفاده از یکی از الگوهای تفاضلی ذکر شده، پیدا کردن برخورد برای تابع چکیده‌ساز LMD4، به طور قطعی امکان‌پذیر بود.

در ادامه از یکی از الگوهای تفاضلی به‌دست آمده برای LMD4 استفاده شد تا تابع چکیده‌سازCMD4، یک نمونه ساده شده دیگر از تابع چکیده‌ساز MD4 تحلیل شود. برای این منظور، احتمال اینکه در تابع چکیده‌ساز CMD4، تابع بولی MAJ مثل تابع بولی XOR عمل کند و احتمال اینکه هیچ رقم نقلی تولید نشود، برای هر مرحله از تابع فشرده‌ساز CMD4 به‌دست آورده شده و با استفاده از این احتمالات پیچیدگی حمله برخورد روی CMD4 (با استفاده از الگوی تفاضلی مورد نظر)، برابر 2^{35} بار محاسبه تابع فشرده‌ساز CMD4 حاصل شد.

در نهایت الگوهای تفاضلی برای تحلیل تابع چکیده‌ساز MD4 بررسی شدند که بنا به دلایلی که توضیح داده شد هیچ یک از این الگوها برای به‌دست آوردن برخورد روی تابع چکیده‌ساز MD4 مناسب نبودند و لازم است تا برای تحلیل تابع چکیده‌ساز MD4 با استفاده از خطی-سازی، روش جستجوی ابتکاری به کار رفته در این مقاله بهبود داده شده و با استفاده از آن الگوهای تفاضلی دیگری پیدا شود.

۷. مراجع

[1]. Biham, E.; Chen, R. “Near-Collisions of SHA-0.”; In Advances in Cryptology, CRYPTO 2004, volume 3152 of LNCS, pages 290-305, Springer Verlag, 2004.

[2]. Brier, E.; Khazaei, S.; Meier, W.; Peyrin, T. “Linearization Framework for Collision Attacks: Application to Cube Hash and MD6.”; ASIACRYPT2009, Springer- Verlage 2009.

[3]. Chabaud, F.; Joux, A. “Differential Collision in SHA-0.”; Advances in Cryptology CRYPTO 98, vol. 1462 of Lecture Notes in Computer Science, pages 56–71. Springer-Verlag, 1998.

[4]. Damgard, I.; “A Design Principle for Hash Functions.”; In Advances in Cryptology, CRYPTO 89, vol. 435 of LNCS, 56-71, Springer- Verlag, 1989.

[5]. Dobbertin, H. “Cryptanalysis of MD4.”; FSE, LNCS 1039, Springer- Verlage, 1996.

[6]. Khazaei, S.; Knellwolf, S.; Meier, W.; Stefan, D. “Improved Linear Differential Attacks on Cube Hash.”; In the Proceedings of AFRICACRYPT 2010, LNCS 6055, 407-418.

[7]. Merkle, R. “One Way Functions and DES.”; In Advances in Cryptology, CRYPTO 89, vol. 435 of LNCS, Springer- Verlag, 1990.

[8]. Mendel, F.; Nad, T. “A Distinguisher for the Compression Function of SIMD-512.”; In Proceedings of INDOCRYPT, LNCS 5922, 219-232, Springer 2009.

[9]. Matusiewicz, K.; Pieprzyk, J. “Finding Good Differential Patterns for Attacks on SHA-1.”; In Coding and Cryptography, WCC’ 2005, vol. 3969 of LNCS, 164–177, Springer 2005.

[10]. Pramstaller, N.; Rechberger, C.; Rijmen, V. “Exploiting Coding Theory for Collision Attacks on SHA-1.”; Cryptography and Coding 2005, LNCS 3796, 78-95, Springer- Verlag, 2005.

[11]. RIPE Consortium, Ripe Integrity Primitive; “Final Report of RACE Integrity Primitve Evaluation (R1040).”; vol. 1007 of LNCS, Springer- Verlag, 1995.

[12]. Rivest, R. L.; “The MD4 Message Digest Algorithm.”; CRYPTO 1990. LNCS, vol. 537, 303–311, Springer, Heidelberg, 1991.

[13]. Rivest, R. “The MD5 Message Digest Algorithm.”; Request for Comments (RFC), 1321, Internet Activities Board, Internet Privacy, Task Force, 1992.

[14]. Rijmen, V.; Oswald, E. “Update on SHA-0.”; CT- RSA 2005, Springer 2005.

[15]. Wang, X.; Lai, X.; Feng, D.; Chen, H.; Yu, X. “Cryptanalysis for Hash Functions MD4 and RIPEMD.”; EUROCRYPT 2005, LNCS, vol. 3494, 1–18. Springer, Heidelberg, 2005.