1st National Conference on Soft Computing of Engineering Science in Industry and Society (ASEIS 2022)
همایش ملی محاسبات نرم علوم مهندسی در صنعت و جامعه

مروری بر کاربرد ریاضیات در کدگذاری

فاطمه راشدی *
استادیار، دانشگاه ولایت، ایرانشهر، f.rashedi@velayat.ac.ir

چکیده. نظریه کدگذاری شاخه ای از ریاضیات می باشد که روش های کنترل خطاهای بوجود آمده در انتقال اطلاعات را بررسی می کند. در واقع امروزه برای تبادل اطلاعات از یک نقطه به نقطه دیگر و یا ذخیره آن‌ها به روی محیط‌های مغناطیسی، اطلاعات را به صورت دنباله ای از نمادها (معمولاً صفر و یک) درآورده و به وسیله دستگاه‌های الکترونیکی ویژه، این دنباله را بین مبدأ و مقصد مبادله می‌کند. جنبه‌های نظری تبادل اطلاعات عمدتاً بر پایه مفاهیم ریاضی و تحت عنوان کدگذاری و کدگشایی مورد بررسی قرار می‌گیرد. امروزه به دلیل نیاز جامعه به خصوص در حوزه مخابرات بیسیم انواع مختلفی از کدها به کار گرفته می‌شوند. در این مقاله به مروری بر کاربرد ریاضیات و ابزارهای جبری در نظریه کدگذاری پرداخته می‌شود.

واژه‌های کلیدی: .کدگذاری، میدان‌های متناهی، هندسه جبری، خم‌های بیضوی.
طبقه‌بندی موضوعی [۲۰۱۰]: 12E20, 94B15, 945B25.

۱. مقدمه

کدگذاری علمی است که ریشه در مخفی‌سازی اطلاعات دارد که در عمل مخفی نگه‌داشتن اطلاعات فردی و جلوگیری از دسترسی افراد غیر مجاز به آن می باشد. در بسیاری از موارد مجبور هستیم برخی از اطلاعات خصوصی خود را به نحوی مخفی نگه داریم. این مخفی کردن در ساده‌ترین حالت پوشاندن موضوع در یک لفافه و قرار دادن آن در جای امن خواهد بود. در مواردی مجبور هستیم پیامی را به صورت مخفی از مکانی به مکانی دیگر منتقل نمائیم در این حالت در لفافه قرار دادن و پنهان نمودن موضوع در یک جای امن کمکی نخواهد کرد بلکه با این موضوع مواجه هستیم که انتقال یک متن از مبدا به مقصد می‌تواند تهدیدی برای افشا شدن موضوع باشد. اولین کدگذاری توسط گایوس یولیوس کایسار[1]، سردار رومي انجام شده است. در این روش کدگذاری که به کد سزار معروف است. کاراکترهاي موجود در متن اولیه براساس ترتیب در حروف الفبا با کاراکترهاي موجود در سه ردیف بعد جایگزین شده و این جایگزيني تا آخرین حروف متن انجام مي شود. با توجه به اینکه در گذشته کدگذاری برای مخفی کردن اطلاعات مورد استفاده قرار می‌گرفت ولی در حال حاضر علاوه بر پنهان سازی اطلاعات برای اقداماتی چون احراز هویت، امضای الکترونیک و . . . نیز به کار می‌رود. طراحی الگوریتم‌های متفاوت کدگذاری که دارای قوت و ضعف متفاوتی هستند و همچنین توسعه و پیشرفت سخت افزار موجب شده است که علم کدگذاری دارای تحول به سزایی شود. کدگذاری به عنوان یک علم از مقاله معروف شانون آغاز شد. او در این مقاله با تعریف کانالي که ورودي آن متن اصلي (متني که مي‌خواهيم کدگذاری شود) و خروجي آن متن کدشده مي‌باشد با تعریف اطلاعات متقابل صفر بین ورودي و خروجي، یک مدل ریاضي مناسب براي کدگذاری ارائه داد. در این تعریف متن اصلي با کلید کدگذاري به گونه‌ای ترکیب مي‌شود که در صورت داشتن متن کدشده و بدون داشتن کلید، نتوان از متن اصلي اطلاعاتي به دست آورد. پس از این مقاله، کدگذاری شروع به پیشرفت نمود. علم کدگذاری را مي‌توان به دو بخش کدگذاري و کدکشایی یا تحلیل کد تقسیم کرد که ارتباط تنگاتنگي بین این دو بخش وجود دارد و پیشرفت در یک بخش منجر به تلاش جهت پیشرفت در بخش دیگر خواهد شد.

*فاطمه راشدی

[1] Gaius Iulius Caesar

الگوريتم‌هاي كدگشايی را مي‌توان به دو دسته عمده تقسيم‌بندي نمود. دسته اول، الگوريتم‌هاي كدگذاری با كليد متقارن هستند كه در آن، دو طرف ارتباط(فرستنده و گيرنده) بايستي بر روي يک پارامتر مخفي به عنوان كليد توافق نمايند و در صورت دست‌يابي گيرنده غير مجاز به آن، باعث شكست الگوريتم كدگذاری و بي‌اثر شدن عمل كدگذاری مي‌گردد. در نتيجه اين كليد به عنوان كليد اصلي الگوريتم كدگذاری، بايد از طريق يک كانال كاملاً امن در اختيار فرستنده و گيرنده قرار گيرد. ساختار اين الگوريتم‌ها از سادگي نسبي برخوردار بوده و درستي و محرمانگي پيام را به طور همزمان تامين مي‌نمايند. اساسي‌ترين مشكل الگوريتم‌هاي رمز متقارن، نحوه توليد و مديريت كليد در الگوريتم‌هاي با تعداد كاربر زياد است. دسته دوم، الگوريتم‌هاي كليد عمومي هستند كه در اواخر دهه ۱۹۷۰ توسط ديفي[۲] و هلمن[۳] ارائه شد كه از خصوصيات عمده آن‌ها مي‌توان به عدم نياز به كانال امن جهت توزيع كليد اشاره كرد. هركدام از دو طرف يک كليد كدگذاری جداگانه براي خود دارند و لازم نيست اين كليد كد، مخفي نگه داشته شود. اما كليد كدگشايي براي هر طرف بايد براي همان طرف معلوم و شخصي بوده و از ديد ديگران مخفي باشد. عمليات كدگذاري و كدگشايي توسط اين دو كليد كه به نام‌هاي كليد عمومي و كليد خصوصي معروف مي‌باشند، انجام مي‌گيرد. اين دو كليد به گونه‌اي انتخاب مي‌شوند كه امكان محاسبه كليد خصوصي از روي كليد عمومي ميسر نباشد مگر آن كه اطلاعات خاصي در اختيار باشد كه تنها خود فرد به آن‌ها دسترسي دارد. در اين الگوريتم‌ها، محرمانگي و درستي پيام را مي‌توان به صورت جداگانه تأمين نمود. گرچه الگوريتم‌هاي كدگذاری كليد عمومي تا حدودي توانسته‌اند مشكل توزيع كليد را حل نمايند اما محدوديت كاربرد آن‌ها و مشكلات پياده‌سازي سخت‌افزاري اين الگوريتم‌ها و سرعت كمتر آن‌ها به دليل حجم محاسبات بالا نسبت به الگوريتم‌هاي رمزنگاري با كليد امن مخصوصا وقتي كه طول متن اصلي زياد باشد، باعث شده كه الگوريتم‌هاي رمزنگاري با كليد امن جايگاه خود را به عنوان نوعی ازالگوريتم‌هاي كدگذاری كارآمد حفظ نمايند. دو كاربرد مهم الگوريتم‌هاي نامتقارن، امضاي ديجيتال داده‌ها و استفاده از آن‌ها در تامين احراز اصالت مي‌باشد.

نظريه كدگذاری كانال يكی از شاخه‌های پركاربرد مخابرات است كه هدف از آن ارسال اطلاعات از طريق يک كانال فيزيكی دارای اغتشاش به گيرنده است. اگر چه اين نظريه در رشته‌های مهندسی برق و كامپيوتر دارای قدمتی طولانی است اما در سال‌های گذشته به دليل توجه خاص به تحليل مباحث تئوری در ساخت و كدگشايی نظر بسياری از رياضيدانان را به خود جلب كرده است. بسياری از شاخه‌های رياضی مانند هندسه، جبر، گراف و تركيبيات در اين زمينه نقش موثر دارند. در اين مقاله به چند مورد از كاربردهای رياضيات در نظريه كدگذاری اشاره می‌شود.

۲. كدگشايی كدهای دوری به كمک حلقه‌ها

تعريف ۱.۲. فرض كنيد كه $A = \{a_1, a_2, \cdots, a_q\}$ يک مجموعه q-عضوی باشد. در اين صورت منظور از يک q-تايی واژه از طول n روی مجموعه A يک دنباله به صورت $w = w_1 w_2 \cdots w_n$ است كه هر $w_i \in A$. به طور معادل w را به صورت $w = (w_1, w_2, \cdots, w_n)$ نشان می‌دهند. منظور از كد q-تايی به طول n روی A، مجموعه ناتهی C متشكل از تمام q-تايی واژه‌هايی است كه دارای طول n باشند. اگر كد روی ميدان $\mathbb{F}_2 = \{0, 1\}$ تعريف شده باشد، آنگاه آن را كد دوبعدی می‌نامند. هر عضو C را يک كدواژه می‌نامند [۴].

تعريف ۲.۲. يک $\mathbb{F}_q$-زيرفضای برداری از $\mathbb{F}_q^n$ را يک كدخطی روی ميدان $\mathbb{F}_q$ می‌نامند. فرض كنيد C يک كدخطی به طول n روی $\mathbb{F}_q$ باشد. در اين صورت دوگان كد C عبارت است از:

$$C = \{u \in \mathbb{F}_q^n \mid uv = 0, \forall v \in C\}.$$

منظور از بعد كد خطی C بعد C به عنوان فضای برداری روی $\mathbb{F}_q$ است [۴].

تعريف ۳.۲. يک كد خطی C از طول n روی ميدان $\mathbb{F}_q$ دوری ناميده می‌شود، هرگاه برای هر $(c_0, c_1, \cdots, c_{n-1}) \in C$، $(c_{n-1}, c_0, c_1, \cdots, c_{n-2}) \in C$ [۴].

تعريف ۴.۲. ماتريس مولد يک كد خطی C ، ماتريسی مانند G است كه سطرهای آن تشكيل يک پايه برای كد C می‌دهند.

يكی از تعميم‌های مهم كد دوری، كد دوری دو بعدی است. به دليل ساختار جبری مناسب كدهای دوری، ماتريس مولد برای كد دوری و دوگان آن به راحتی به دست می‌آيد و در نتيجه كدگشايی آن‌ها به آسانی امكان‌پذير است. اما ماتريس مولد

[2] Diffie
[3] Hellman

برای کد و دوگان یک کد دوری دو بعدی (یکی از تعمیم‌های مهم کد دوری) به دست نیامده است. برای استفاده از خواص جبری در کدهای دوری از تناظر زیر استفاده می‌شود:

$$f : \mathbb{F}_{q^n} \longrightarrow \mathbb{F}_{q^n}/\langle x^n - 1\rangle$$
$$(c_0, c_1, \cdots, c_{n-1}) \longmapsto c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}.$$

در [۲] به بررسی ساختار برخی از کدهای دوری دو بعدی پرداخته می‌شود. در واقع کدهای دوری دوبعدی متناظر با ایده‌ال‌های حلقه $\mathbb{F}[x,y]/\langle x^s - 1, y^{2k} - 1\rangle$ بررسی شده است و همچنین ماتریس مولد برای کد خطی و دوگان کد خطی به دست آمده است. در کدهای دوری روی یک خانواده نامتناهی از حلقه‌ها تعریف می‌شوند. خواص کلی کدهای دوری بر روی این حلقه‌ها مورد مطالعه قرار می‌گیرد، به‌ویژه کدهای دوری نابدیهی با یک مولد مشخص می‌شون [۱].

۳. کدهای تصحیح کننده خطاها

قسمت وسیعی از اطلاعات که در کدگذاری مبادله می‌شود، در قالب اعداد نشان داده می‌شوند. در واقع اطلاعات به صورت دنباله‌ای از اعداد که به طور فیزیکی متناظر با علائم الکترونیکی یا علائم دیگر کدگذاری شده‌اند. در مجموع اطلاعات به شکل دنباله‌ای از ارقام دودویی یعنی اعداد ۰ و ۱ کدگذاری شده‌اند. یک مسأله بزرگ در کدگذاری خطاها می‌باشند. یکی از راه‌های مقابله با این خطاها، کدهای تصحیح کننده خطاها می‌باشند. یک کد تصحیح کننده خوب باید دارای بازدهی بالایی باشد و قابلیت خوبی در کشف و تصحیح خطاها داشته باشد و کدگشایی را ساده و سریع کند. در واقع کدهای تصحیح کننده خطاها با طولانی کردن پیام‌ها کدهایی می‌سازند که تا حد امکان خطاها را کشف و تصحیح کنند و کدگشایی آن‌ها آسان باشد.

۴. کاربرد میدان‌های متناهی در کدهای تصحیح کننده خطاها

در سال ۱۹۸۴ کلود شانن [۴] ریاضیدان آمریکایی کد بهینه‌ای برای تصحیح خطاها معرفی کرد، اما روشی برای ساخت آن‌ها پیدا نکرد. نظریه شانن وجود کدهای تصحیح کننده خوبی را اثبات کرد. علاوه بر این کدهای همینگ که قابلیت متوسطی داشتند توسط ریچارد همینگ[۵] در سال ۱۹۵۰ مطرح شده بود. بعد از این متخصصین سعی کردند تا کدهای تصحیح کننده معرفی شده توسط شانن را مورد مطالعه قرار دهند. آن‌ها برای مطالعه کدهای تصحیح کننده از نظریه میدان‌های متناهی که توسط اوراریست گاما[۶] هنگام مطالعه معادلات جبری مطرح شده است استفاده کردند. به این ترتیب به کمک جبر مجرد در ارتباط با نظریه میدان‌های متناهی کدهای تصحیح کننده خطای موثری ساخته شد [۳].

۵. کاربرد هندسه جبری در کدگذاری

متخصصین کدهای تصحیح کننده از هندسه جبری که بخش وسیعی از ریاضیات کنونی است استفاده کرده‌اند. هندسه جبری به بررسی اشیاء هندسی از جمله خم‌ها، رویه‌ها و ... که توسط معادلات جبری تعریف می‌شود، می‌پردازد. علاوه بر این منحنی‌های تعریف شده روی میدان‌های متناهی نیز مطالعه می‌شود به این صورت که در معادلات جبری نمایشگر آن‌ها کمیت‌هایی نظیر x و y دیگر اعداد دلخواه نیستند بلکه منحصر به یک میدان متناهی خاص هستند. با استفاده از این منحنی‌ها و جبر وابسته به مختصات نقاط آن‌ها خانواده جدیدی از کدهای تصحیح کننده و کدهای هندسی ساخته شده است. این کدها باعث شدند که نتایج جدیدی مربوط به کدهای دوبعدی به دست آید و کدهایی با قابلیت بیشتری از کدهای شانن ساخته شوند. در مقابل تحلیل کدهای هندسی، ریاضیدان‌ها را به بررسی دقیق‌تر در مورد نقاط یک منحنی جبری که بر روی یک میدان تعریف شده است، هدایت کرده است.

[4] Claude Shannon
[5] Hamming Richard
[6] Evariste Galois

۶. کاربرد خم‌های بیضوی در کدگذاری

در سال ۱۹۷۰ کدگذاری به روش RSA اختراع شد. در این روش از دو کلید استفاده می‌شود که یک کلید کدگذاری عمومی است که همه می‌توانند آن را بشناسند و یک کلید کدگشا است که محرمانه باقی می‌ماند. این روش متکی بر این اساس است که می‌توان اعداد اول بزرگی با صر رقم یا هزار رقم ساخت ولی یافتن عوامل اول p و q از روی یک عدد بزرگ $N = p \times q$ بسیار مشکل است. در واقع شناخت عدد N شناخت کلید عمومی کدگذاری است و شناخت p و q شناخت کلید محرمانه کدگشا است. در روش RSA نظریه اعداد در سطح پیشرفته‌ای دخالت دارد. در سال‌های اخیر روش‌های خوبی در کدگذاری ظاهر شده‌اند که بر اصولی نزدیک به اصول RSA متکی هستند. یکی از این روش‌ها همان روش لگاریتم گسسته [۷] می‌باشد. این روش سبب شد که روش‌های دیگری که بر ویژگی‌های خم‌های بیضوی بنا می‌شوند، به وجود آید. این خم‌ها به شکل بیضی نیستند بلکه خم‌هایی هستند که مطالعه آن‌ها در قرن ۱۹ برای حل مسأله پیچیده محیط بیضی مطرح شد. این خم‌ها ویژگی‌های جانبی دارند. درواقع به کمک یک ساختمان هندسی می‌توان به شکلی عمل جمع را بین نقاط خم بیضوی تعریف کرد. به طور کلی، خم‌های بیضوی اشیائی هستند که دارای ویژگی‌های حسابی هستند که می‌توانند در کدگذاری استفاده شوند [۵].

۷. نتیجه‌گیری

نظریه کدگذاری کانال یکی از شاخه‌های پرکاربرد مخابرات است که هدف از آن ارسال اطلاعات از طریق یک کانال فیزیکی دارای اغتشاش به گیرنده است. اگر چه این نظریه در رشته‌های مهندسی برق و کامپیوتر دارای قدمتی طولانی است اما در سال‌های گذشته به دلیل توجه خاص به تحلیل مباحث تئوری در ساخت و کدگشایی نظر بسیاری از ریاضیدانان را به خود جلب کرده است. بسیاری از شاخه‌های ریاضی مانند هندسه، جبر، گراف و ترکیبیات در این زمینه نقش موثر دارند. کدگذاری در طول دهه‌های اخیر شاهد پیشرفت‌هایی بوده و به دانشی پیچیده تبدیل شده است. پیشرفت‌های کدگذاری نتیجه کار متخصصینی با آموزه‌های سطح بالای در ریاضیات است. در این مقاله به چند مورد از کاربردهای ریاضیات در نظریه کدگذاری اشاره می‌شود. در واقع به کدگشایی کدهای دوری به کمک حلقه‌ها اشاره می‌شود. علاوه براین به کاربرد میدان‌های متناهی در کدهای تصحیح کننده خطاها و کاربرد هندسه جبری، نظریه اعداد و خم‌های بیضوی در کدگذاری اشاره می‌شود.

مراجع

1. S.T. Dougherty, S. Karadeniz, B. Yildiz, *Cyclic codes over R_k, Designs, Codes and Cryptography*, 63 (2012), 113–126.
2. Z. Sepasdar, K. Khashyarmanesh, *Characterizations of some two-dimensional cyclic codes correspond to the ideals of* $\mathbb{F}[x,y]/\langle x^s-1, y^{2^k}-1\rangle$, Finite fields and their applications 41 (2016), 97–112.
3. G. Lachaud, *Communiquer Sans Erreurs: Les Codes Correcteurs*, L'explosion des mathematiques, SMF et SMAI, Paris, (2002), 84–87.
4. S. Ling, Ch. Xing, *Coding Theory A First Course*, Cambridge University Press, (2004).
5. J. L. Nicolas, *Cryptage et Decryptage: Communiquer en Toute Securite*, L'explosion Des Mathematiques, SMF et SMAI, Paris, (2002), 15–18.

[7] Logarihme Discret