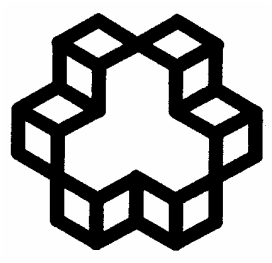


تاسیس ۱۳۰۷

دانشگاه صنعتی خواجه نصیرالدین طوسی

دانشکده مهندسی برق

پایان نامه کارشناسی ارشد در رشته مهندسی برق- مخابرات (سیستم)

عنوان

شبیه‌سازی کدهای LDPC و کاربرد آن در سیستم ADSL

نگارش :

محمد باقر نظافتی

استاد راهنما :

دکتر محمود احمدیان عطاری

مهر ۸۲

تقدیر و تشکر

پس از سپاس از خداوند مهربان بر خود لازم می‌دانم که صمیمانه‌ترین تشکرات خود را از همه دوستانی که در انجام این پژوهش مرا یاری کرده‌اند به جا آورم. ابتدا از استاد راهنمای خویش آقای دکتر احمدیان تشکر می‌کنم که در مدت انجام این پژوهش راهنما و جوابگوی سوالات من بوده‌اند.

از آقای فرشید مجیدفر که پشتیبان من بوده‌اند کمال تشکر را دارم. همچنین از آقای دیانت و بهروزی که راهنمای بنده در انتخاب موضوع پژوهش بوده‌اند و آقای دکتر مک‌کی که جوابگوی تعدادی از سوالات بنده بوده‌اند تشکر و قدردانی می‌نمایم.

از اساتید ممتحن آقای دکتر جمالی و آقای دکتر کلانتری، همکاران دانشگاه صنعتی مالک اشتر و مرکز تحقیقات مخابرات تشکر می‌نمایم.

از کلیه همکاران آموزش، تحصیلات تکمیلی و ریاست دانشکده برق نیز تشکر می‌نمایم.

در پایان از پدر و مادر فداکار و سایر اعضای خانواده‌ام که در همه مراحل زندگی پشتیبان من بوده‌اند کمال تشکر را دارم.

تقدیم به مولای متقیان امیرالمومنین(ع)، پدر و مادر عزیزم

و دوست مهربانم آقای فرشید مجیدفر

چکیده

در این نوشتار کدهای LDPC ساختار گالاگر و مک‌کی شبیه‌سازی می‌شود و تاثیر طول بلوک کد و وزن ستونهای ماتریس بررسی درستی[1] در عملکرد کدها بررسی می‌شود. همچنین نشان داده می‌شود که این دو ساختار، عملکرد مشابهی دارند.

کاربرد نوع خاصی از کدهای LDPC در مودمهای ADSL بررسی می‌شود و نشان داده می‌شود که این کدها در ترکیب با مدولاسیون باینری و چندسطحه عملکرد خوبی دارند و دارای شیب خطا نیستند. به همین دلیل این کدها در کاربردهایی که محدودیت پهنای باند وجود دارد، می‌تواند مناسب باشد. همچنین عملکرد و پیچیدگی این کد با کد TCM و توربو مقایسه می‌شود.

[1] Parity check matrix

فصل اول : مقدمه و تعاریف

مقدمه

کدهای LDPC[1] نوعی کد بلوکی هستند که برای تصحیح خطای کانال بکار می‌روند. این کدها برای اولین بار توسط گالاگر در سال ۱۹۶۰ معرفی شدند[1]. گالاگر خواص مهمی را برای این کدها اثبات کرد. او اثبات کرد که با افزایش طول بلوک کد، احتمال خطا بصورت نمایی کاهش می‌یابد و حداقل فاصله کد بصورت خطی افزایش می‌یابد.

کدهای LDPC به علت نیاز به حافظه زیاد برای کدگذاری و پیچیدگی کدبرداری، در آن زمان مورد توجه قرار نگرفت. در سال ۱۹۹۶ مک‌کی و نیل این کدها را دوباره کشف کردند [2] و نشان دادند که کدهای LDPC جزء کدهای بسیار خوب هستند[3].

مک‌کی الگوریتم کدبرداری Sum-Product را برای کدبرداری این کدها بکار برد و نشان داد که این الگوریتم نتایج بسیار خوبی را بدنبال دارد. مک‌کی و دیوی نوع غیربایزی کدهای LDPC را نیز معرفی کردند [5] و بهبود عملکرد کدهای LDPC غیربایزی را نسبت به کدهای مشابه بایزی نشان دادند.

Luby کدهایی را معرفی کرد که ماتریس بررسی درستی آنها ناهمسان بود [7]. این کدها که کدهای نامنظم نامیده می‌شوند، عملکرد بهتری دارند.

عملکرد بسیار خوب کدهای LDPC موجب شده است که برای کاربردهای مختلف مانند ضبط مغناطیسی [8]، ضبط نوری [9]، CDMA [43] و مخابرات سیمی[10]،[11] پیشنهاد شوند.

در بعضی از سیستمهای مخابراتی سیمی مانند مودم ADSL، جهت بهبود عملکرد سیستم از کدینگ استفاده می‌شود. در استاندارد آمریکایی ANSI برای ADSL از کد RS و یا کد الحاقی مرکب از کد RS و کد TCM، Wei استفاده می‌شود[12].

---

[1] Low Density Parity check Code

اخیراً الفتریو[1] نوعی کد LDPC که ساختار معینی دارد برای ADSL پیشنهاد کرده است. کد پیشنهاد شده می‌تواند جایگزین کد کانولوشنال شود و با افزایش پیچیدگی قابل قبولی سبب بهبود عملکرد سیستم شود[10]،[11].

ساختار ادامه این نوشتار بصورت زیر است.

در فصل دوم کدهای LDPC با ساختار گالاگر و مک‌کی معرفی می‌شوند و خواص فاصله‌ای و الگوریتم طراحی کد و نتایج شبیه‌سازی آنها توسط کدبردار Sum-Product ارایه می‌شود.

در فصل سوم ابتدا الگوریتم کدبرداری Sum-Product وسپس الگوریتم کدبرداری Sum-Product تعمیم یافته برای کدبرداری کدهای غیرباینری معرفی می‌شوند. همچنین در این فصل الگوریتم Sum-Product مورد تحلیل قرار می‌گیرد و با الگوریتم کدبرداری توربو مقایسه می‌شود.

در فصل چهارم انواع مودمها معرفی می‌شوند و نرخ مبادله آنها، مدولاسیون و کدهای بکاررفته در آنها بیان می‌شود.

در فصل پنجم مودم ADSL با جزئیات بیشتری بررسی می‌شود. در این فصل مدولاسیون، نحوه تخصیص بیت به زیرکانالها و کدهای تصحیح خطای بکاررفته در ADSL بررسی می‌شوند.

در فصل ششم کاربرد کدهای LDPC معین در مودم ADSL بررسی می‌شود. در این فصل نحوه طراحی و کدگذاری کدهای LDPC معین بیان می‌شود و سپس نتایج شبیه‌سازی آنها با مدولاسیونهای باینری و چندسطحه ارایه می‌شود و نشان داده می‌شود که منحنی خطای این کدها در ترکیب با مدولاسیونهای چندسطحه مشابه منحنی خطای آنها در ترکیب با مدولاسیونهای باینری است. در پایان عملکرد و پیچیدگی این کد با کد پیشنهاد شده در استاندارد ADSL و توربوکد مقایسه می‌شود.

[1] Eleftheriou

تعاریف

بردار تنک[1]: برداری که اغلب عناصر آن صفر باشد.

وزن بردار: تعداد عناصر غیر صفر بردار را گویند.

همپوشانی دو بردار: تعداد مکانهایی که هر دو بردار عنصر غیرصفر داشته باشند.

چگالی ماتریس: نسبت عناصر غیرصفر به کل تعداد عناصر ماتریس را گویند.

ماتریس کم چگال[2] ***(ماتریس تنک)***: ماتریسی که چگالی آن کوچکتر از ۰/۵ باشد.

ماتریس بسیار تنک[3]: ماتریسی که با افزایش ابعاد آن، چگالی ماتریس به سمت صفر میل کند. (مانند ماتریسی که وزن سطری و ستونی ثابت داشته باشد)

ماتریس منظم[4]: ماتریسی که همه سطرهای آن با هم و همه ستونهای آن با هم وزن یکسانی داشته باشد. کدی را که ماتریس بررسی درستی آن منظم باشد، کد منظم گویند.

ماتریس نامنظم[5]: ماتریسی که تنوع وزنی سطر و ستونهای آن زیاد باشد. کدی که ماتریس بررسی درستی آن نامنظم باشد، کد نامنظم گویند.

حداقل فاصله کد: حداقل فاصله کد D است اگر و تنها اگر (D-1) ستون ماتریس بررسی درستی کد مستقل خطی باشند و D ستون از ماتریس وجود داشته باشند که وابسته خطی باشند.

کدهای عملی: کدهایی که پیچیدگی کدگذار و کدبردار آنها بگونه‌ای باشد که بصورت عملی قابل پیاده‌سازی باشند.

---

[1] Sparse vector
[2] Low density martix
[3] Very sparse matrix
[4] Regular matrix
[5] Irregular matrix

کدهای خوب: کدهایی که حداکثر نرخ آنها عدد مثبتی است که از ظرفیت کانال کوچکتر است. این کدها می‌توانند احتمال خطا را به دلخواه کوچک کنند. کد الحاقی جزء کدهای عملی و خوب به شمار می‌رود.

کدهای بسیار خوب: کدهایی که اگر نرخ آنها از ظرفیت کانال کمتر باشد، می‌توانند احتمال خطا را به دلخواه کاهش دهند. شانون وجود چنین کدهایی را پیشبینی کرده است. با افزایش حدطول کدهای کانولوشنال می‌توان به ظرفیت کانال نزدیک شد ولی با افزایش حدطول کد پیچیدگی کدبردار ویتربی بصورت نمایی افزایش می‌یابد. بنابراین کدهای کانولوشنال جزء کدهای بسیار خوب ولی غیرعملی در حدطولهای بزرگ هستند.

کدهای بد: کدهایی که احتمال خطا را تنها در صورتی به اندازه دلخواه کاهش می‌دهند که نرخ آنها تا صفر کاهش یابد.

آنتروپی باینری: آنتروپی باینری برای کانال BSC بصورت زیر تعریف می‌شود[3].

$$H_2(p) = p\log_2\left(\frac{1}{p}\right) + (1-p)\log_2\left(\frac{1}{1-p}\right) \qquad (۱-۱)$$

p در رابطه (۱-۱) احتمال خطا در کانال BSC می‌باشد.

متوسط آنتروپی: بردار X بطول N و تابع توزیع $P(X)$ دارای متوسط آنتروپی H_X است، اگر به ازای هر $\varepsilon > 0$ و $\eta > 0$ عددی مانند N' وجود داشته باشد که به ازای هر $N > N'$ رابطه زیر برقرار باشد[3].

$$P\left(\left|\frac{1}{N}\log_2\frac{1}{P(X)} - H_X\right| > \eta\right) < \varepsilon \qquad (۲-۱)$$

کدبردار typical set: کدبرداری است که جستجوی بردار کلمه‌کد X بطول N را به جای جستجو در فضای $\{0,1\}^N$ در فضای محدود T که بصورت زیر تعریف می‌شود، انجام می‌دهد[3].

$$T = \left\{ X \in \{0,1\}^N : \left| \frac{1}{N} \log_2 \frac{1}{P(X)} - H_X \le \eta \right| \right\} \qquad (۳-۱)$$

$P(X)$ و H_X در رابطه (۱-۳) بترتیب توزیع X و متوسط آنتروپی X است و η عدد کوچکی است که به دلخواه انتخاب می‌شود.

خطای قابل تشخیص: الگوی خطایی که در معادله بررسی درستی صدق نکند.

خطای غیر قابل تشخیص: الگوی خطایی که در معادله بررسی درستی صدق کند.

فصل دوم: معرفی کدهای LDPC

۲-۱ کدهای تصحیح خطا

کدهای تصحیح خطا برای غلبه بر خطای انتقال کانال بکار می روند. در این کدها با افزودن تعدادی بیت افزونی به رشته بیت ارسالی، بیتهای خطا تصحیح می شود. شانون در سال ۱۹٤۸ تئوری کدینگ کانال نویزی را بیان کرد و پارامتری به نام ظرفیت را برای کانالهای نویزی تعریف کرد[4]. شانون در این تئوری اثبات کرد که اگر نرخ ارسال اطلاعات از ظرفیت کانال کمتر باشد، با افزایش طول بلوک کد می‌توان احتمال خطا را تا حد دلخواه کاهش داد.

این تئوری تنها وجود چنین کدی را اثبات می کند ولی نحوه تولید این کد را نشان نمی دهد. بنابراین از نظر تئوری، دستیابی به هر احتمال خطای دلخواه امکان‌پذیر است ولی در عمل به علت افزایش پیچیدگی و تاخیر کدگذار و کدبردار دستیابی به هر احتمال خطای دلخواه عملی نیست. به همین علت طراحی کدهایی که با پیچیدگی قابل قبول به ظرفیت کانال نزدیک باشند، حائز اهمیت است.

۲-۱-۱ کدهای خطی

کدهای تصحیح خطای خطی توسط ماتریس بررسی درستی H و یا ماتریس مولد G توصیف می‌شود. ماتریس G دارای K سطر و N ستون است و کلمه کد V بصورت زیر از حاصلضرب ماتریس مولد G و بردار پیام U حاصل می شود.

$$V = UG \qquad (۱-۲)$$

ماتریس مولد را در صورتی سیستماتیک گویند که بصورت $G = [I_K \quad P]_{K\times N}$ باشد که I_K ماتریس واحد به ابعاد $K\times K$ و P ماتریس باینری به ابعاد $K\times(N-K)$ است.

فرم سیستماتیک ماتریس بررسی درستی بصورت $H = \left[P^T \quad I_{N-K}\right]_{(N-K)\times N}$ تعریف می شود و دارای خواص زیر است.

$$GH^T = 0 \qquad (۲-۲)$$

بردار نویز W توسط کانال به بردار V افزوده می شود و در نتیجه بردار دریافتی X بصورت زیر در گیرنده بدست می آید.

$$X = UG + W \qquad (۳-۲)$$

کدبردار با دریافت بردار X و داشتن فرضیاتی درباره خصوصیات کانال بردار U را بازیابی می‌کند. کدبردار بهینه[1] بردار U را چنان بازیابی می کند که احتمال پسین[2] زیر حداکثر شود.

$$P(U \mid X, G) = \frac{P(X \mid U, G)P(U)}{P(X \mid G)} \qquad (۴-۲)$$

از حاصلضرب ماتریس H در طرفین معادله (۲-۳) و استفاده از رابطه (۲-۱) معادله زیر حاصل می شود.

$$Z = H \times X^T = H \times W^T \qquad (۵-۲)$$

بردار Z را در رابطه فوق، بردار سندرم گویند. بنابراین کدبرداری به مساله یافتن بردار W که در رابطه زیر صدق کند تقلیل می یابد.

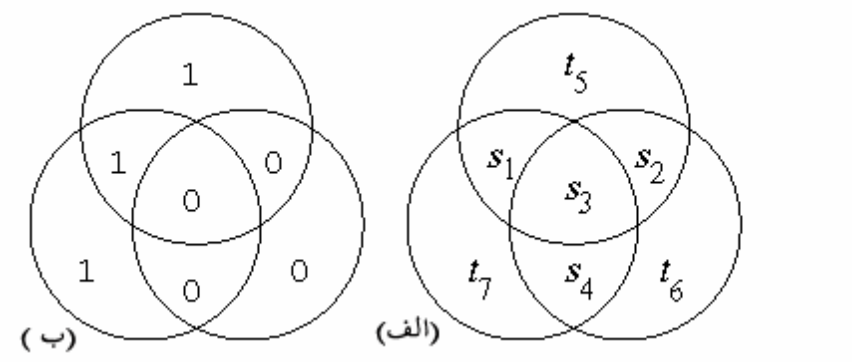
$$Z = H \times W^T (\mathrm{mod}\, 2) \qquad (٦-۲)$$

۲-۱-۲ نمایش کدها به وسیله گراف

اغلب کدها را می توان توسط گرافی که به گراف کد معروف است نمایش داد. به عنوان مثال کد همینگ (٤ , ۷) را در نظر بگیرید. نحوه کد گذاری این کد در شکل (۲-۱-الف) نمایش داده شده است. در این شکل ۷ بیت ارسالی در سه دایره هم پوشان توزیع شده است. چهار بیت t_4 , t_3 , t_2 ,t_1 به بیت‌های

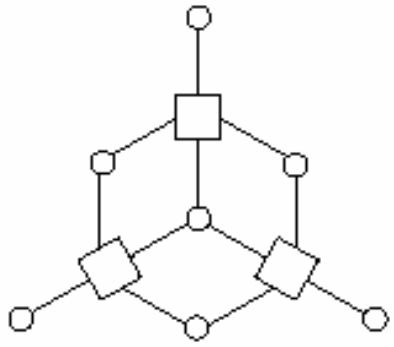
[1] Optimum decoder
[2] Posteriori probability

پیام S_1, S_2, S_3, S_4 نسبت داده می شود و بیت های پریتی t_5, t_6, t_7 چنان محاسبه می شود که پریتی دایره ها زوج باشد. بیت های پریتی با فرض S=1000 در شکل (۲-۱-ب) محاسبه شده اند[13].



شکل(۲-۱): نمایش گرافیکی کد همینگ (۷ , ٤)

اگر به جای بیتهای ارسالی گره متغیر و به جای دایره‌ها گره چک قرار دهیم به گرافی می‌رسیم که آن را گراف دو بخشی[1] کد همینگ (۷ , ٤) می نامند. این گراف در شکل (۲-۲) نمایش داده شده است. در این شکل گره‌های متغیر بصورت دایره و گره‌های چک بصورت مربع نمایش داده شده است[13].



شکل (۲-۲): گراف کد همینگ (۷ , ٤)

گراف دو بخشی به گرافی گفته می شود که گره‌های آن را بتوان به دو دسته متغیر و چک تقسیم کرد بگونه‌ای که هر شاخه گرهی از یک گروه را به گرهی از گروه دیگر متصل کند و هیچ شاخه‌ای دو گره از یک گروه را به هم متصل نکند[19]. این گراف را Influence Diagram، Causal Network ، Bayesian Network و Belief Network نیز می‌نامند.

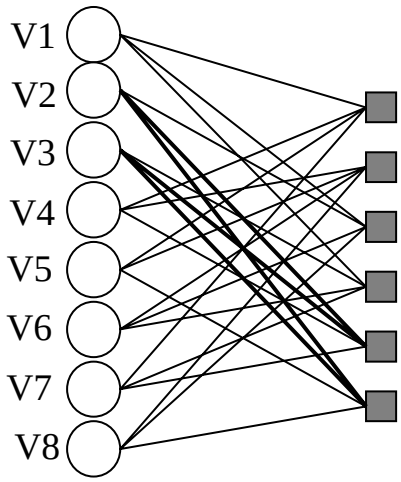
[1] Bipartite graph

گراف دو بخشی متناظر ماتریس بررسی درستی کد، H است. گره‌های متغیر متناظر با ستونهای ماتریس و گره‌های چک متناظر با سطرهای ماتریس است. اگر H دارای M سطر و N ستون باشد، گراف دو بخشی آن دارای M گره چک و N گره متغیر خواهد بود و به ازای هر m و n که در رابطه زیر صدق می کند، شاخه‌ای گره متغیر n را به گره چک m متصل می کند.

$$H[m][n]=1 \quad m=1,2,...,M \quad n=1,2,...,N \qquad (۷-۲)$$

تعداد شاخه‌های متصل به هر گره را درجه گره می نامند. درجه گره‌های چک با وزن سطرهای H و درجه گره‌های متغیر با وزن ستونهای H برابر است. در شکل (۲-۳) مثالی از ماتریس بررسی درستی و گراف دو بخشی آن نمایش داده شده است. در این شکل گره‌های سمت چپ که با دایره نشان داده شده است گره متغیر و گره‌های سمت راست که با مربع نشان داده شده است، گره‌های چک هستند[43].

$$\begin{pmatrix} 1&0&0&1&1&0&1&0 \\ 0&0&0&1&1&1&0&1 \\ 1&1&0&0&0&1&0&1 \\ 1&0&1&0&0&1&1&0 \\ 0&\mathbf{1}&\mathbf{1}&1&0&0&1&0 \\ 0&\mathbf{1}&\mathbf{1}&0&1&0&0&1 \end{pmatrix}$$



شکل (۲-۳): ماتریس بررسی درستی و گراف دو بخشی معادل آن

حلقه به مسیر بسته‌ای گفته می‌شود که از یک گره شروع می شود و پس از پیمودن شاخه‌هایی دوباره به آن گره ختم می شود. حلقه‌های بطول 2L (طول حلقه همیشه زوج است) بصورت زیر بیان می‌شود.

$$V_1,C_1 \;\; C_1,V_2 \;\; V_2,C_3 \;... \; C_L,V_{L+1}=V_1$$

با توجه به ساختار گراف دو بخشی، کوچکترین حلقه دارای طول ۴ می باشد. طول کوچکترین حلقه گراف را گرث[1] می نامند. حلقه های به طول ۴ از وجود دو تلاقی در ستونهای ماتریس حاصل می‌شود. نمونه ای از حلقه های به طول ۴ با خطوط پررنگ در گراف شکل(۲-۳) و با یک‌های پر رنگ در ماتریس بررسی درستی آن نشان داده شده است.

۲-۲ کدهای LDPC با ساختار گالاگر

کدهای LDPC با ماتریس بررسی درستی تنک[2] که اغلب عناصر آن صفر و تعداد بسیار کمی یک است معرفی می شوند. نوعی از این کدها که به کدهای گالاگر منظم معروف است از نخستین کدهای LDPC به شمار می رود. کدهای ساختار گالاگر با پارامترهای (N, t, r) تعریف می شود که در آن N طول بلوک کد، t وزن ستونهای ماتریس بررسی درستی و r وزن سطرهای آن است[1].

ماتریس بررسی درستی کدهای ساختار گالاگر(N, t, r) دارای N ستون است که تعداد یک‌ها در هر ستون برابر t و در هر سطر برابر r است و بقیه عناصر ماتریس صفر است. بنابراین ماتریس بررسی درستی دارای Nt/r سطر است و نرخ کد بصورت $R \geq 1-t/r$ می باشد.

شکل (۲-۴) ماتریس بررسی درستی کدی را به ازای $r = 4 , t = 3 , N = 20$ نشان می‌دهد. این ماتریس به t زیر ماتریس تقسیم می شود که هرکدام تنها دارای یک ۱ در هر ستون است. یک‌ها در ستون اول به حالت نزولی قرار دارند، بگونه‌ای که در سطر i ام بین ستونهای $(i-1)r+1$ و ir قرار می گیرند. زیرماتریسهای دیگر این ماتریس تنها جایگشت زیر ماتریس اول است. خانواده کدهای گالاگر (N , t , r) به کدهایی گفته می شود که از جایگشت ستونهای هر یک از زیرماتریسهای دوم به بعد حاصل می شود.

[1] Girth
[2] Sparse

1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0
0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0
0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1
1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0
0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0
0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1

شکل (۲-۴): ماتریس بررسی درستی کد گالاگر منظم با پارامترهای r = 4, t = 3, N = 20

در بخش بعد نشان داده می شود که حداقل فاصله کد به ازای r , t ثابت و $t \geq 3$ بطور خطی با طول بلوک افزایش می یابد ولی در صورتیکه t = 2 باشد، حداقل فاصله کد (N , 2 , r) با لگاریتم N افزایش می یابد.

۲-۲-۱ خواص فاصله ای کدهای گالاگر منظم

تابع فاصله‌ای[1] کدهای گالاگر با $N(\ell)$ نشان داده می‌شود و بصورت تعداد کلمات کد با وزن ℓ تعریف می شود. از خواص گروهی کدهای بررسی درستی می توان نتیجه گرفت که $N(\ell)$ تعداد کلمات کد با فاصله ℓ از کلمه کد مشخص نیز می باشد. حداقل فاصله کد که با D نشان داده می شود، بصورت کوچکترین مقدار $\ell > 0$ که به ازای آن $N(\ell) \neq 0$ باشد تعریف می شود. واضح است که D باید در هر کدی به اندازه کافی بزرگ باشد و $N(\ell)$ به ازای $\ell > D$ تا حد ممکن کوچک باشد.

محاسبه تابع فاصله کد و یا حداقل فاصله کد برای N های بزرگ عملاً غیر ممکن است، به همین علت معمولاً متوسط تابع فاصله کد روی خانواده‌ای از کدها بررسی می شود.

[1] Distance function

قبل از بررسی خواص فاصله‌ای کدهای گالاگر منظم، به معرفی خانواده دیگری از کدهای بررسی درستی که به کدهای بررسی درستی هم احتمال[1] معروف است می پردازیم. این کدها توسط Elias در بررسی باندهای کدهای تصادفی برای کدهای بررسی درستی معرفی شده است[20]. اگر طول بلوک کد N و نرخ کد R باشد، ماتریس بررسی درستی کد تصادفی هم احتمال دارای N(1-R) سطر و N ستون است و مکانهای ماتریس با رقمهای باینری هم احتمال و مستقل از هم پر می شود.

قضایای ۱،۲ و قضایای ۳،۴،۵ که در زیر بیان می شوند، بترتیب خواص فاصله ای کدهای هم احتمال و کدهای گالاگر را بیان می کنند.

قضیه ۱ : اگر $\overline{N(\ell)}$ متوسط تعداد کلمات کد با وزن ℓ روی مجموعه ای از کدهای هم احتمال با طول بلوک N و نرخ R باشد، $\overline{N(\ell)}$ به ازای $\ell > 0$ در نامساوی زیر صدق می کند[1].

$$\overline{N(\ell)} = \binom{N}{\ell} 2^{-N(1-R)} \le [2\pi N\lambda(1-\lambda)]^{-\frac{1}{2}} \exp\big(N[H(\lambda)-(1-R)Ln2]\big)$$
$$\lambda = \frac{\ell}{N}\ , \qquad H(\lambda) = \lambda Ln\frac{1}{\lambda} + (1-\lambda)Ln\frac{1}{1-\lambda} \qquad (۸-۲)$$

قضیه ۲ : اگر $P(D \le \delta N)$ تابع توزیع حداقل فاصله کدهای بررسی درستی هم احتمال با طول بلوک N و نرخ R باشد، $P(D \le \delta N)$ به ازای $\delta < \frac{1}{2}$ و δN صحیح در نامساوی زیر صدق می کند[1].

$$P(D \le \delta N) \le \frac{1}{1-2\delta}\sqrt{\frac{1-\delta}{2\pi N\delta}} \exp N[H(\delta)-(1-R)Ln2]$$
$$P(D \le \delta N) \le 1 \qquad (۹-۲)$$

باند فوق با افزایش N به تابع پله در نقطه δ_0 میل می کند که $\delta_0 < \frac{1}{2}$ و $H(\delta_0) = (1-R)Ln2$ است. نسبت δ_0 برحسب نرخ کد در شکل (۲-۵) نشان داده شده است.

[1] Equiprobable parity check code