نظریه اطلاعات کلاسیک_ بخش دوم

وحید کریمی پور_ دانشکده فیزیک _ دانشگاه صنعتی شریف

۲۵ دی ۱۳۹۸

۱ مقدمه

می خواهیم رشته هایی از 0 و 1 را از یک کانال مخابراتی که دارای نوفه است مخابره کنیم. اما این کانال تحت تاثیر نوفه [۱] قرار دارد و هرعلامت 0 یا 1 کهارسال می کنیم بااحتمال q تبدیل به 1 یا 0 شده و با احتمال $1-q$ دست نخورده باقی می ماند. می خواهیم کاری کنیم که گیرنده پیام ها همچنان بتواند پیام های صحیح را ازپیام های دریافت شده استخراج کند. راهی که برای تعیین وتصحیح خطا باید پیش بگیریم آن است عنصر تکرار را به نوعی وارد پیام های خود کنیم. به عنوان مثال می توانیم بجای یک 0 سه تا 0 وبجای یک 1 سه تا 1 مخابره کنیم، یعنی از کد تکرار سه تایی استفاده کنیم

$$0 \longrightarrow 000 \qquad 1 \longrightarrow 111 \tag{۱}$$

واز گیرنده بخواهیم که با استفاده از قانون اکثریت تصمیم بگیرد که یک رشته سه تایی که دریافت کرده است درواقع چه رشته ای بوده است . درواقع احتمال خطا که قبلاً برابربود با q اینک کمترشده است، زیرااحتمال وقوع خطا اینک برابراست با احتمال وقوع دو برگردان و یا سه برگردان درکد سه تایی که اولی برابر است با $3q^2(1-q)$ ودومی برابراست با q^3. درنتیجه احتمال وقوع خطا برابرخواهد بود با

$$P_E = q^3 + 3q^2(1-q) \tag{۲}$$

[۱] Noise

که برای q های کوچک ازمرتبه q^2 است. البته بهایی برای این کاهش خطا پرداخت کرده ایم و آن این است که نرخ مخابره اطلاعات را پایین آورده ایم و از سه بیت برای مخابره یک بیت استفاده کرده ایم. دراین حالت نرخ مخابره اطلاعات یعنی R برابراست با $\frac{1}{3}$. اما این نرخ کم است. اما هنوز احتمال خطا وجود دارد اگر چه احتمال وقوع آن را نسبت به قبل کمتر کرده ایم؟ آیا می توانیم احتمال وقوع خطا را پایین تر بیاوریم؟ پاسخ اش مثبت است. بجای کد تکرار سه تایی می توانیم کد تکرار ۵ تایی یعنی کد زیر استفاده کنیم:

$$0 \longrightarrow 00000 \qquad 1 \longrightarrow 11111. \tag{۳}$$

فاصله این دو رشته از هم برابر با ۵ است و احتمال وقوع خطا برابر است با احتمال اینکه سه تا بیشتر از این بیت ها دچار خطا شوند . این خطا برابر است با

$$P_E = \binom{5}{3} q^3(1-q)^2 + \binom{5}{4} q^4(1-q) + q^5 \sim 10q^3 \tag{۴}$$

که به مراتب از میزان خطای قبلی کمتر است. اما برای کاهش خطا مجبور شده ایم که نرخ مخابره اطلاعات را بازهم پایین بیاوریم و آن را به $R = \frac{1}{5}$ برسانیم. آیا می توانیم کدی به کار ببریم که نرخ اش از این مقدار بالاتر باشد؟ برای این کار می توانیم به جای اینکه تک تک علائم 0 و 1 را کد کنیم بلوک های دوتایی از این علایم را کد کنیم. مثلا چیزی شبیه به این:

$$\begin{aligned} 00 &\longrightarrow 00000 \\ 01 &\longrightarrow 11100 \\ 10 &\longrightarrow 00111 \\ 11 &\longrightarrow 11011. \end{aligned} \tag{۵}$$

در این کد نیز فاصله هر دو رشته از یک دیگر برابر با سه است و بنابراین احتمال خطا از مرتبه q^3 است ولی نرخ مخابره پیام کمی بهتر شده است و از $\frac{1}{3}$ به $\frac{2}{5}$ رسیده است. به نظر می رسد که با کد کردن رشته های دوتایی به جای تک تک حروف الفبا توانسته ایم ضمن اینکه احتمال خطا را کاهش می دهیم نرخ را نیز کمی بالا ببریم. آیا می توانیم بهتر از این عمل کنیم؟ مثلا چطور است که بلوک های چهارتایی را کد کنیم، آنهم توسط کد هامینگ یعنی کد $[4, 7, 3]$. این کد فاصله ۳ دارد و چهار بیت را در ۷ بیت کد می کند. به این ترتیب نرخ مخابره را از $\frac{2}{5}$ نیز کمی بالاتر می برد و به $\frac{4}{7}$ می رساند، ضمن اینکه احتمال خطا را از همان مرتبه q^3 نگاه می دارد. ولی خب بازهم احتمال خطا وجود دارد و ما به هیچ وجه نمی خواهیم مخابره ای با احتمال خطا انجام دهیم چرا که از بین میلیون ها بیت مخابره شده ده ها هزار بیت کاملا به اشتباه مخابره خواهند شد. آنچه که ما می خواهیم این است که احتمال خطا را به سمت صفر میل دهیم. اما همانطور که در کد تکرار دیدیم به نظر می رسد که برای

پایین آوردن احتمال خطا به صفر ما مجبوریم نرخ مخابره را نیز به صفر میل دهیم. آیا راهی برای برون رفت از این اشکال وجود دارد؟ پاسخ اش مثبت است. راه اش را در مثال های بالا دیدیم. راه اش این است که بلوک های بزرگ تر را کد کنیم. یعنی k بیت را به n بیت کد کنیم. در این صورت نرخ مخابره برابر خواهد بود با $R = \frac{k}{n}$ و ممکن است که با بزرگ تر گرفتن هر چه بیشتر k و n این نسبت به سمت یک مقدار غیر صفر میل کند و بسیار مهم تر از آن شاید بتوانیم همزمان احتمال وقوع خطا را نیز به صفر برسانیم. در این صورت حد $\lim_{n\longrightarrow\infty} \frac{k}{n}$ را ظرفیت کانال می نامیم. این در واقع محتوی قضیه دوم شانون یا قضیه کانال نوفه دار شانون [۲] است. این قضیه بیان می کند که برای هر کانال مخابراتی یک ظرفیت مثل C وجود دارد که می بایست با توجه به مشخصات کانال آن را محاسبه کرد. در این صورت همواره می توان با نرخی کمتر از این مقدار ظرفیت مخابره اطلاعات انجام داد بدون اینکه مرتکب هیچ نوع خطایی بشویم. به عبارت دقیق تر به ازای هر دو عدد کوچک ϵ و δ همواره یک نوع کدگذاری خاص مثل C^* و یک عدد صحیح مثل $n*$ وجود دارد که اگر طول رشته ها را بزرگ تر از n^* بگیریم در این صورت به ازای آن کد گذاری خاص هر دو شرط زیر توامان برآورده می شوند:

$$R < C - \delta \qquad P_E(x) < \epsilon \qquad \forall x, \tag{۶}$$

که در آن $P_E(x)$ احتمال خطا در مخابره رشته x است.

برای اینکه این قضیه و اثبات آن را بفهمیم نخست

حالت ساده ای را درنظرمی گیریم که منبع X پیام هایی را ارسال می کند که درآنهااحتمال وقوع 0 و 1 یکسان باشد. به عبارت دیگر آنتروپی منبع دراین مثال برابراست با 1. فرض کنید که پیام خود را دررشته های n بیتی ارسال کنیم . مجموعه تمام رشته های n بیتی دارای 2^n رشته است. نقاط این فضا نقاط یک شبکه ابرمکعبی n بعدی را پرکرده اند. اگر همه این نقاط را به عنوان کدرشته های خود به کاربریم به آسانی درطول عبوراز کانال یک کد رشته به یکی از کد رشته های همسایه اش یا کد رشته های نزدیکش تبدیل می شود و گیرنده واقعاً نمی فهمد که چه کد رشته ای برایش ارسال شده است. بنابراین راه مقابله با خطا آن است که سعی کنیم که کدرشته ها را بافاصله کافی از یکدیگر انتخاب کنیم تااحتمال وقوع خطا پایین بیاید. درواقع وقتی که یک کدرشته را ارسال می کنیم حول آن یک کره به اندازه کافی بزرگ رسم می کنیم و هروقت نقطه ای درون این کره دریافت کنیم آن را به عنوان کد رشته ای که درمرکز آن قراردارد تعبیر وخطاهای بوجود آمده را تصحیح می کنیم. این کاربه طور طبیعی نرخ را پایین می آورد زیرازهمه نقاط شبکه استفاده نمی کنیم. حال می پرسیم که کد رشته ها را با چه فاصله ای انتخاب کنیم. دراین جا می بایست بین دو عامل متضاد موازنه ایجاد کنیم. اگر بخواهیم خطارا پایین بیاوریم می بایست فاصله کد رشته ها را از یکدیگر زیاد کنیم. ازطرفی این کارنرخ را پایین می آورد. خوب بیایید ببینیم به طور متوسط یک کد رشته به چند کدرشته نزدیک دیگر ممکن است تبدیل شود؟

[۲] Shannon Noisy Channel Theorem

در یک رشته n حرفی هر حرف با احتمال q ممکن است که برگردانده شود. بنابراین تعداد حرف های برگردانده شده به طور متوسط برابراست با nq. این تعداد حرف می تواند به $\left(\begin{array}{c} n \\ nq \end{array} \right) \approx 2^{nH(q)}$ طریق در رشته n حرفی قرارگیرند. درنتیجه تعداد رشته های نزدیک به این رشته که ممکن است ازخطاهای بوجود آمده دراین رشته ایجاد شوند برابراست با $2^{nH(q)}$. دراین جا نیز ایده اصلی این است که اگر چه کل خطاهای ممکن در یک رشته n تایی برابراست با 2^n اما تعداد خطاهای متعارف یا نمونه از این مقدار کمتر است. در حقیقت اگر وقوع یک خطا را با بیت 1 و عدم وقوع آن را با بیت 0 نشان دهیم آنگاه تمام خطاهای ممکن تبدیل می شوند به رشته هایی از صفر و یک که تعدادشان برابراست با 2^n ولی نکته در این است که ما تنها می بایست به خطاهای نمونه یا $typical$ فکر کنیم که تعدادشان برابر است با $2^{nH(q)}$ که کمتر از مقدار کل است. در اینجا تابع $H(q)$ همان آنتروپی شانون برای خطا ست که برابراست با

$$H(q) = -q \log_2 q - (1-q) \log_2 (1-q). \qquad (۷)$$

تمامی قضایایی که در درس گذشته قاره تعداد و احتمال رشته های نمونه یاد گرفتیم در مورد خطاهای نمونه نیز صادق است.

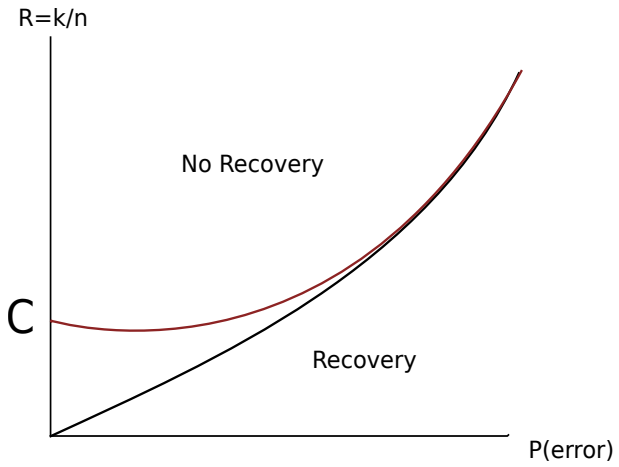
درنتیجه یک نحوه کد گذاری خوب آن است که حول هررشته ناحیه ای درنظربگیریم که به طور متوسط تعداد رشته های داخل آن درحدود $2^{nH(q)}$ باشد. حال اگر نرخ مبادله برابربا $R = \frac{k}{n}$ باشد معنایش این است که تعداد کل کدرشته ها ها برابراست با $2^k = 2^{nR}$. چون تعدادکل رشته های ممکن برابراست با 2^n به این نتیجه می رسیم که می بایست شرط زیربرقراراباشد:

$$2^{nR} 2^{nH} \leq 2^n, \qquad (۸)$$

یا

$$R \leq 1 - H(q) =: C. \qquad (۹)$$

کمیت $C := 1 - H(q)$ ظرفیت این کانال نامیده می شود واین رابطه بیان می کند که برای پرهیزازخطا نر خ مبادله اطلاعات حتما باید ازظرفیت کانال کمترباشد. در واقع آنچه که نشان داده ایم این است بر خلاف تصور اولیه می توان منحنی نرخ بر حسب خطا را از سقوط به نقطه صفر نجات داد و با توجه به این نکته که خطاهای نمونه خیلی کمتر از تعداد کل رشته های ممکن است می توانیم یک نوع کدکردن به کار ببریم که ضمن آنکه نرخ خطا را در حد مجانبی به صفر می رساند ولی نرخ مبادله اطلاعات را به سمت یک مقدار حدی میل می دهد. هر گونه مبادله اطلاعات با کمتر از این نرخ و بدون خطا ممکن است اما هرگونه تجاوز از این نرخ باعث می شود که هیچ گونه اطلاعاتی قابل بازیابی نباشد.



شکل ۱ : رابطه نرخ مبادله اطلاعات و نرخ خطا و مفهوم ظرفیت.

حال بایک سوال اساسی مواجه می شویم و آن اینکه آیا واقعا نوعی کدگذاری وجود دارد که بتوان بااستفاده از آن نرخ مبادله اطلاعات رادرحد مجانبی $n \longrightarrow \infty$ به حداکثر ممکن یعنی به C رساند؟ پاسخ این سوال مثبت است . در واقع نشان می دهیم که حتما یک کد وجود دارد که می تواند نرخ مبادله بدون خطای اطلاعات را هرچقدر که می خواهیم به ظرفیت C نزدیک کند. به عبارت دیگر نشان می دهیم که حتما یک کد وجود دارد که با به کار بردن آن می توانیم نرخ مبادله خطا را هرچقدر که می خواهیم بهC نزدیک کرده و نرخ خطا را هرچقدر که می خواهیم کوچک کنیم. دقت کنید که در اینجا مسأله ما ساختن صریح این کد نیست بلکه می خواهیم نشان دهیم که یک چنین کدی وجود دارد. برای اثبات به ترتیب زیر پیش می رویم. فضای تمام های n بیتی را در نظر می گیریم. این فضا دارای 2^n نقطه است. حال به طور تصادفی 2^k نقطه را درون این فضا در نظر می گیریم. (درست مثل پرتاب کردن دارت به یک صفحه.) این حرف به این معناست که با احتمال یکنواخت این نقاط را انتخاب می کنیم. بنابراین نحوه کدکردن ما به همین سادگی است، یعنی به جای تعریف صریح کد (مثلا نوشتن کد خطی با یک ماتریس مولد یا ماتریس پاریته و نظایرآن) تنها2^k نقطه را به طور تصادفی در این فضا اختیار می کنیم و این نقاط را به عنوان کد رشته های خود به کار می بریم. به این ترتیب می توانیم k بیت را در n بیت کد کنیم. این کد را یک کد تصادفی می نامیم و آن را باC_1 نشان می دهیم. دقت کنید که ممکن است بعضی از رشته ها کد به هم نزدیک انتخاب شوند که در این صورت احتمال رخ دادن خطا برای این کد رشته ها زیاد است. در عوض در بعضی موارد فاصله یک کد رشته می تواند از بقیه خیلی زیاد باشد که در این صورت احتمال رخ دادن خطا یعنی تبدیل این کد رشته به یک کد رشته دیگر خیلی کم است. هر بار که این نقاط را انتخاب می کنیم یک کد جدید به طور تصادفی مشخص می شود. این کد ها را با C_1 ، C_2 و نظایر آن نشان می دهیم.

■ قضیه: به ازای هر مقدار کوچک ϵ و هر مقدار کوچک δ حتما یک کد $\mathcal{C}^*$ وجود دارد به قسمی که :

احتمال خطا برای تمام رشته ها از ϵ کمتر باشد و نرخ مبادله پیام نیز هر چقدر که می خواهیم به ظرفیت نزدیک باشد به این معنا که شرط $C - R < \delta$ برقرار باشد.

برای اثبات این قضیه به ترتیب زیر عمل می کنیم. هر رشته ای که دریافت می کنیم به دورش یک کره به شعاع $(H(q) + \delta)$ رسم می کنیم. اصطلاحاً این به این معناست که مجموعه تمام نقاطی را در نظر می گیریم که دارای $2^{n(H(q)+\delta)}$ در بر داشته باشند. هرگاه یک کد رشته درون این کره قرار داشت می گوییم رشته ای که ارسال شده است همان نقطه بوده است و رشته دریافت شده را به همان رشته اصلاح می کنیم، شکل ۲. خطای بازگشایی وقتی رخ می دهد که رشته دیگری از جای دیگری درون کره مورد نظر ما بیفتد، شکل (۳). بجای اینکه به یک کد مشخص نگاه کنیم مجموعه تمام کدهای تصادفی را در نظر می گیریم و احتمال چنین خطایی را به طور متوسط حساب می کنیم. این متوسط روی تمام کدهای تصادفی و هم چنین روی تمام رشته ها در هر کدام از کدهاست. این احتمال خطا را با $\langle \overline{P} \rangle$ را نشان می دهیم. این نماد در واقع نشان دهنده دو نوع متوسطی است که در بالا به آن اشاره شد. به طور دقیق تر برای یک کد مشخص مثل $\mathcal{C}$ داریم

$$\overline{P^{\mathcal{C}}} = \frac{1}{2^k} \sum_{x=1}^{2^k} P^{\mathcal{C}}(x) \qquad (۱۰)$$

که در آن $P^{\mathcal{C}}(x)$ احتمال این است که در کد $\mathcal{C}$ ما در بازگشایی رشته x دچار خطا شده باشیم. هم چنین علامت براکت برای محاسبه متوسط روی همه کدهای تصادفی اختیار شده است یعنی به ازای هر کمیت مثل O

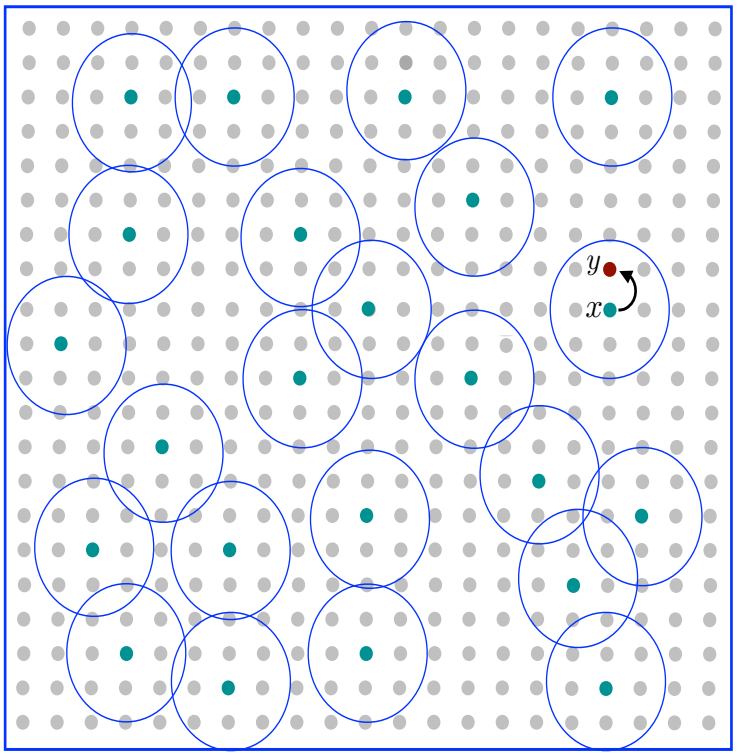
$$\langle O \rangle = \frac{1}{\mathcal{Z}} \sum_{\mathcal{C}} O(\mathcal{C}), \qquad (۱۱)$$

که در آن $\mathcal{Z}$ تعداد کدهای تصادفی است. حال می توانیم $\langle \overline{P} \rangle$ را حساب کنیم. این احتمال را بدون توجه به جزییات کدها می توان حساب کرد. می خواهیم ببینیم احتمال اینکه نقطه دیگری از یک کره دیگر به درون یک کره مورد نظر ما بیفتد چقدر است؟ در اثر خطا هر نقطه سبز رنگی می تواند شروع به حرکت کند و در جای دیگری از این فضا که شامل 2^n نقطه است بنشیند. اگر یک کره مشخص را حول نقطه ای مثل x در نظر بگیریم ممکن است نقطه ای از یک کره نزدیک حرکت کرده و درون این کره قرار گرفته باشد یا اینکه نقطه ای از یک کره دوردست دچار خطای زیاد شده و به درون این کره رسیده باشد. احتمال اینکه یک رشته (یک نقطه سبزرنگ) در اثر خطا (یا ولگشت) به درون این کره افتاده باشد برابر است با:

$$\frac{2^{n(H(q)+\delta)}}{2^n}. \qquad (۱۲)$$

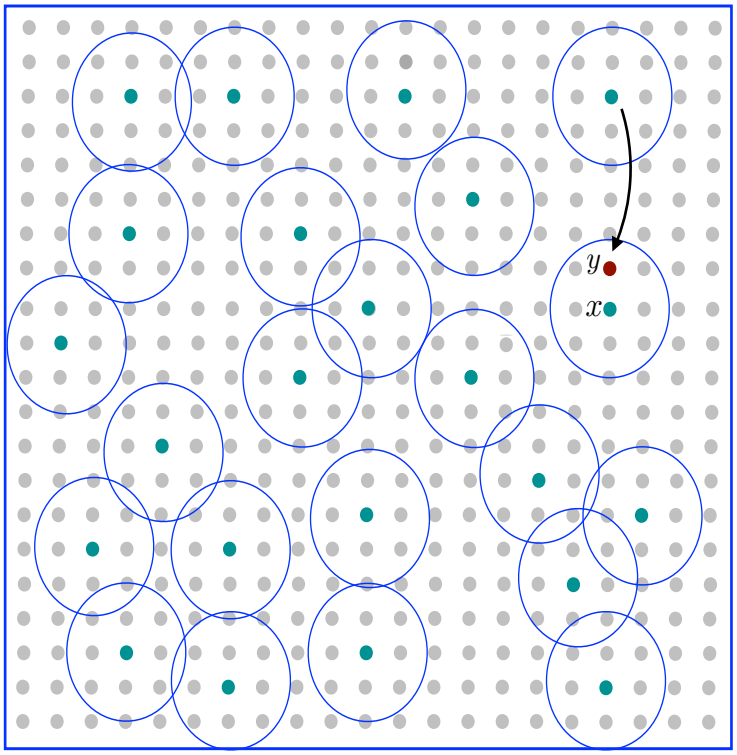
تعداد کل رشته ها یا نقاطی که می توانسته اند دچار خطا شده باشند برابر است با $2^k = 2^{nR}$. بنابراین احتمال متوسط خطا برابر ا ست با:

$$\langle \overline{P} \rangle = 2^{nR} \times \frac{2^{n(H(q)+\delta)}}{2^n} = 2^{n(R-C+\delta)}. \qquad (۱۳)$$



شکل ۲ : هر رشته کد شده با رنگ سبز نشان داده شده است. به دور هر رشته یک کره رسم می کنیم. در اثر خطا ممکن است رشته ای که در مرکز یک کره است جابجا شود و به نقطه ای دیگر (به رنگ قرمز) درون همان کره منتقل شود. در انتهای کانال اگر چنین نقطه ای دریافت کنیم آن را به همان نقطه مرکز کره برمی گردانیم چرا که به احتمال زیاد نقطه قرمز همان نقطه سبز بوده که در اثر خطای کانال جابجا شده است. ،

آنچه که این استدلال مبتنی بر نسبت حجم ها را موجه می کند این است که در طرف چپ، دو نوع متوسط گرفته شده است یکی متوسط روی



شکل ۳: خطا در بازگشایی رشته ها وقتی رخ می دهد که یک نقطه از یک کره دیگر دراثر خطا به درون کره مورد نظر ما می افتد. در این صورت این رشته را به رشته ای تصحیح می کنیم که ربطی به رشته فرستاده شده ندارد.

تمام رشته ها در هر کد و دوم متوسط روی تمام کدهای تصادفی. این رابطه نشان می دهد که با انتخاب مناسب n می توانیم نرخ مبادله اطلاعات یعنی R را هرچقدر که می توانیم به C نزدیک کنیم و درعین حال خطای متوسط را هرچقدر که می خواهیم کم کنیم.

■ تمرین:اگر بخواهیم که متوسط احتمال خطا کمتر از 0.001 شود و$R \leq C - 0.01$باشد طول کد رشته ها حداقل چقدر باید باشد؟ اگر بخواهیم با همین نرخ اطلاعات را مبادله کنیم ولی متوسط احتمال خطا را به زیر 0.00001 برسانیم حداقل طول کد رشته ها چقدر خواهد بود؟

استدلال بالا نشان می دهد که احتمال خطا را به طور متوسط می توان از هر مقداری که بخواهیم کمتر کنیم. ولی این خطا در واقع یک خطای متوسط دوگانه است به این معنا که یک بار روی تمامی کد رشته های درون یک کد متوسط گرفته ایم و باردیگر روی تمام کدهای تصادفی $C_1, C_2, \cdots$ متوسط گرفته ایم. به طور دقیق تر ثابت کرده ایم که می توان با انتخاب طول مناسب برای کد رشته ها نرخ مبادله اطلاعات را هرچقدر که می خواهیم به ظرفیت کانال نزدیک کنیم آنچنان که متوسط احتمال خطا از هر مقداری که می خواهیم کمتر باشد یعنی

$$\langle \overline{P} \rangle \leq \epsilon. \qquad (۱۴)$$

حال از رابطه ۱۴ نتیجه می شود که حتما یک کد تصادفی مثل C وجود دارد که برای آن رابطه زیر برقرار است:

$$\overline{P^{\mathcal{C}}} \leq \epsilon. \qquad (۱۵)$$

اما این عبارت متوسط خطا را روی تمام کدرشته ها نشان می دهد. ممکن است برای بعضی از کد رشته ها خطا بیشتر از ϵ و برای بعضی دیگر کمتر از ϵباشد. این کد البته مطلوب ما نیست چرا که ما کدی می خواهیم که احتمال خطا برای تمام کدرشته هایش از ϵ کمتر باشد. برای رسیدن به این کد تنها کافی است که کد را کمی خلوت کنیم و بعضی از کد رشته ها را حذف کنیم. تنها یک نگرانی وجود دارد و آن اینکه با خلوت کردن کد و دور ریختن این نوع رشته ها تعداد خیلی کمی کد رشته در کد باقی بماند و نرخ به شدت پایین بیاید یا حتی به صفر برسد. نشان می دهیم که چنین نیست. در واقع نشان می دهیم که تنها با دور ریختن کمتر از نیمی از کد رشته ها (یا رشته ها) می توانیم به کدی برسیم که خطا برای تک تک کدـرشته ها از 2ϵ کمتر باشد و چون 2ϵ نیز یک مقدار بی نهایت کوچک است به هدف خود دست پیدا می کنیم. برای این کار بازهم از لم اول چبیشف استفاده می کنیم.

احتمال خطای مربوط به کد رشته اول را با $P^{\mathcal{C}}(x_1)$ ، خطای مربوط به کد رشته دوم را با $P^{\mathcal{C}}(x_2)$ نشان می دهیم و همینطور تا آخر. در این صورت می دانیم که متوسط این خطا ها از ϵ کمتر است. یعنی

$$\overline{P^{\mathcal{C}}} = \frac{1}{2^{nR}} \sum_{x=1}^{2^{nR}} P^{\mathcal{C}}(x) \leq \epsilon. \qquad (۱۶)$$

اگر x را به عنوان یک متغیر تصادفی ببینیم و $P^{\mathcal{C}}(x)$ را به عنوان تابعی از متغیرتصادفی x در نظر بگیریم، می توانیم لم چبیشف را برای آن به کار ببریم و بنویسیم:

$$Prob(P^{\mathcal{C}}(x) \geq 2\epsilon) \leq \frac{\overline{P^{\mathcal{C}}}}{2\epsilon} \leq \frac{\epsilon}{2\epsilon} = \frac{1}{2}. \qquad (۱۷)$$

معنای این حرف این است که کمتر از نصف تعداد کل کد رشته ها احتمال خطایشان از 2ϵ بیشتر است و اگر این تعداد از کد رشته ها را دور بریزیم بقیه همه احتمال خطایشان از 2ϵ کمتر است و این همان چیزی است که می خواستیم ثابت کنیم. بنابراین با دور ریختن کمتر از نیمی از کد رشته ها به کدی مثل C^* می رسیم که احتمال خطا برای تمام کد رشته هایش از 2ϵ کمتر است. کد جدیدی که ساخته ایم تعداد کد رشته هایش برابر است با:

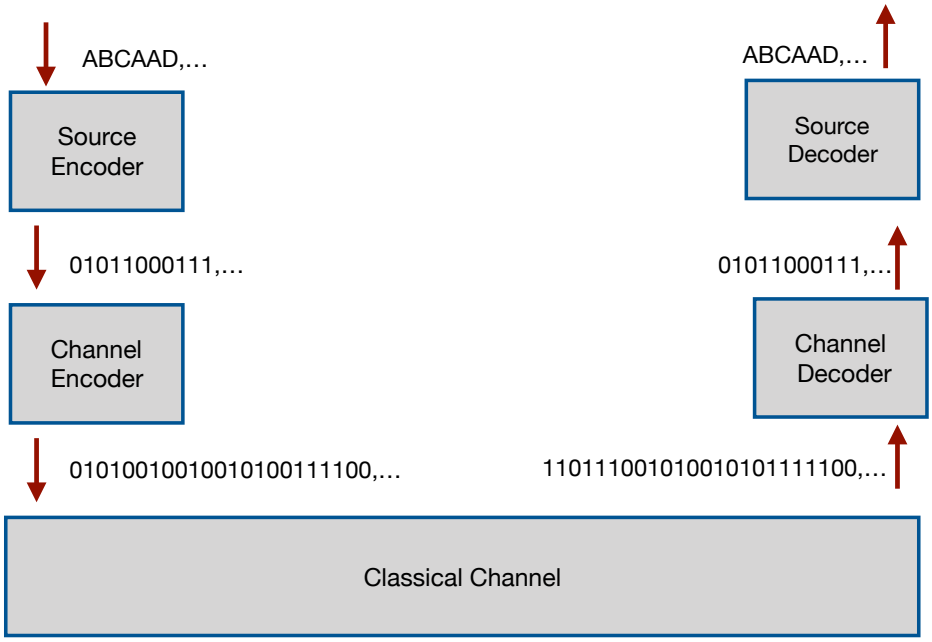
$$\frac{2^{nR}}{2} = 2^{n(R-\frac{1}{n})} \quad (۱۸)$$

و این به این معناست که نرخ جدید در حد n های بزرگ با نرخ قبلی برابر است. بنابراین ثابت کردیم که یک کد وجود دارد که نرخ مبادله اطلاعات را به حد شانون یعنی ظرفیت کانال می رساند و در عین حال خطا را هم هر چقدر که بخواهیم کم می کند.

۲ تعریف ظرفیت درحالت کلی

بهتر است یکبار دیگر به کل یک کانال مخابراتی کلاسیک از ابتدا تا انتها نگاه کنیم، شکل (۴). پیام اولیه نخست وارد یک کدکننده منبع [۳] می شود. کار این کدکننده این است که متن را مطابق با آنچه که در درس قبل دیدیم، فشرده کند. در بهترین حالت این فشردگی قرار است به حد شانون برسد اگرچه ممکن است در عمل این کار انجام نشود. دقت کنید که این کد کننده الزاما حروف را به تعداد بیت های یکسان کد نمی کند. همانطور که در یک مثال الفبای چهارتایی دیدیم این کدکننده ممکن است حرف A را به یک بیت و حرف B را به دو حرف کد کند. به این ترتیب خروجی این کدکننده رشته بسیار بلندی از علایم صفر و یک است. این رشته بسیار بلند قبل از ورود به کانال می بایست بازهم تبدیل به یک کد شود، این بار نه برای فشرده کردن بلکه برای مقابله با خطاهایی که ممکن است در کانال صورت بگیرد. به این ترتیب رشته بسیار بلند به رشته های k تایی تقسیم شده و هر رشته مطابق با یک کد تصحیح خطا به یک رشته n تایی کد می شود. این قسمت توسط کدکننده کانال [۴] صورت می گیرد. در انتهای کانال رشته های دریافت شده نخست وارد یک کدگشای کانال [۵] می شوند تا تصحیح خطا روی آنها انجام شده و سپس وارد کدگشای منبع [۶] می شوند تا تبدیل به الفبای اولیه شوند. مجموعه کلیه رشته های k تایی ممکن برابر است با 2^k. در نتیجه مجموعه رشته های n تایی نیز برابر است با 2^k. این مجموعه را با $X^{(n)}$ نمایش می دهیم. اعضای این مجموعه را با $x^{(n)}$ نشان می دهیم. این رشته ها بسته به نوع کدگذاری که به کار رفته است آنتروپی مشخصی خواهند داشت. این آنتروپی را با $H(X^{(n)})$ نشان می دهیم. هر رشته $x^{(n)}$ که به کانال وارد می شود به صورت یک رشته $y^{(n)}$ خارج می شود. این احتمال شرطی را با $P(y^{(n)}|x^{(n)})$ نمایش می دهیم. مجموعه تمام رشته های

[۳] Source Encoder
[۴] Channel Encoder
[۵] Channel Decoder
[۶] Source Decoder



شکل ۴: نمای کلی یک کانال کلاسیک.

خروجی را با $Y^{(n)}$ نشان می دهیم. آنتروپی این مجموعه را نیز با $H(Y^{(n)})$ نشان می دهیم. تعداد نمونه های این رشته ها با تقریب بسیار خوب برابر است با $2^{H(Y^{(n)})}$. هر رشته $x^{(n)} \in X^{(n)}$ که در نظر بگیریم در اثر خطا می تواند به رشته های متفاوتی از $Y^{(n)}$ تبدیل شود. احتمال این که چنین رشته ای به $y^{(n)}$ تبدیل شود برابر است با $P(y^{(n)}|x^{(n)})$. به ازای یک $x^{(n)}$ مشخص مجموعه تمام $y^{(n)}$ هایی که حاصل می شوند را می توان نشان دهنده یک متغیر تصادفی با احتمال $P(y^{(n)}|x^{(n)})$ در نظر گرفت. تعداد رشته های نمونه از این نوع به طور متوسط برابر است با $2^{H(Y^{(n)}|X^{(n)})}$ ، که در آن متوسط روی همه رشته های $x^{(n)}$ گرفته شده است. بنابراین روش کد کردن چنین است که به گرد هر رشته $x^{(n)}$ کره ای با شعاع $H(Y^{(n)}|X^{(n)})$ رسم می کنیم. این کره دارای $2^{H(Y^{(n)}|X^{(n)})}$ رشته است. حال شرط این که کره های مختلف با هم تداخل نکنند این است که

$$2^k \times 2^{H(Y^{(n)}|X^{(n)})} \leq 2^{H(Y^{(n)})} \qquad (۱۹)$$

که از آن نتیجه می گیریم

$$\frac{k}{n} \leq \frac{H(Y^{(n)}|X^{(n)}) - H(Y^{(n)})}{n} = \frac{I(Y^{(n)} : X^{(n)})}{n} \qquad (۲۰)$$

و چون همه قضایایی که در مورد رشته های نمونه داشته ایم در حد $n \longrightarrow \infty$ صادق اند، به تعریف زیر برای ظرفیت کانال می رسیم:

$$C := Max_{P(X)} \lim_{n \longrightarrow \infty} \frac{I(Y^{(n)} : X^{(n)})}{n}. \qquad (۲۱)$$

توجه به این نکته مهم است که روی تمامی توابع توزیع ورودی مقدار بیشینه را حساب کرده ایم.