

25.

Přenosová a modulační rychlost, Informační rychlost, Shannon – Hartleyův zákon, Propustnost kanálů, Entropie

Tuto otázku jsem pojal trochu dvěma směry. V podtitulu „To, co by vám mělo stačit ke státnicím“ jsem psal jen ty nezákladnější a zároveň nejpodstatnější věci, které byly k danému tématu probrány v daných předmětech.

Na internetu jsem našel také nějaké materiály a ty jsem umístil do podtitulu „Pro toho, kdo se chce dovědět víc“. Myslím, že není nutné si toto číst, dal jsem to sem jen pro to, že některé věci jsou tam docela dobře vysvětleny z trochu jiné strany a tak se to dá i dobře pochopit. Takže pokud nebudete mít co dělat s volným časem, můžete si to přečíst... :-) Zároveň upozorňuji, že některé informace se v této „nepovinné“ sekci opakují a také některé informace nejsou zrovna aktuální (to ale pro hlubší pochopení látky není podstatné).

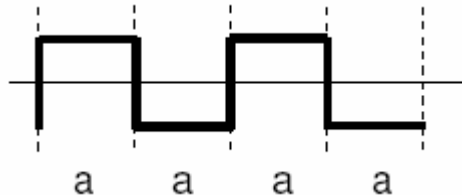
Tak jen doufám, že vám bude tato otázka k užitku a případné nesrovnalosti nebo nejasnosti můžete zasílat na můj školní mail: velisl1@fel.cvut.cz

Modulační rychlost

To, co by vám mělo stačit ke státnicím:

Asi nezákladnější definice modulační rychlosti (v_m) je, že **modulační rychlost udává počet signálových prvků (chceme-li symbolů) přenesených za sekundu**. Jinak řečeno je to rychlost, se kterou se mění jednotlivé stavy nosné. Tato rychlost se vyjadřuje v jednotkách *Baudech* [Bd] a je rovna převrácené hodnotě doby trvání (a) signálového prvku:

$$v_m = \frac{1}{a} \text{ [Bd, s}^{-1}\text{]}$$



Pozn.: Čím větší modulační rychlost, tím menší trvání (a) jednotlivých signálových prvků a tím naopak rychlejší změny mezi jednotlivými stavy.

Další důležitou věcí, kterou je dobré si uvědomit je, že **modulační rychlost závisí na šířce použitého pásma**. Z toho tedy plyne, že nám určuje požadavky na mezní kmitočet f_m určitého přenosového kanálu.

Ze vztahu

$$v_m \leq 2 \cdot f_m$$

plyne, že modulační rychlost nemůžeme u daného kanálu zvyšovat nad hodnotu $2f_m$.

Pro toho, kdo se chce dovědět víc:

Zastavme se nyní u jedné velmi zajímavé otázky: jak často si můžeme dovolit měnit průběh přenášeného signálu při přenosech v přeloženém pásmu? Neboli: jak rychle (resp. jak často) můžeme „hýbat“ s parametry přenášeného signálu, které v rámci používané modulace měníme? Omezujícím kritériem je samozřejmě to, aby příjemce na druhé straně dokázal tyto změny dostatečně spolehlivě detekovat. Formálně se přitom ptáme na tzv. modulační rychlost (modulation speed, někdy též: Baud rate), která vyjadřuje počet změn za sekundu, a měří se v jednotkách zvaných Baud (Bd).

Odpověď na tuto otázku nám objasní tzv. Nyquistovo kritérium. To nám říká, že ani při snaze o maximální „vyždímání“ přenášeného signálu nemá smysl jej vzorkovat rychleji než dvakrát za každou jeho periodu, neboli rychlostí číselně dvojnásobnou oproti dostupné šířce přenosového pásma. Odsud pak můžeme stanovit jednoduchý vzoreček pro maximální možnou modulační rychlost:

$$v_m = 2 \cdot \text{šířka pásma } (f_m)$$

Povšimněme si hned jedné zajímavé a důležité vlastnosti: maximální dosažitelná modulační rychlost závisí pouze na dostupné šířce přenosového pásma, a nikoli na konkrétní použité modulaci.

Přenosová rychlost

To, co by vám mělo stačit ke státnicím:

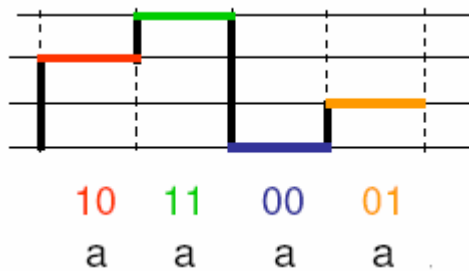
Číselný signál může obecně nabývat různých hodnot (stavů). U dvoustavových modulací nese každý signálový prvek (symbol) informaci o jediném bitu. Například u čtyřstavových modulací už nese však informaci o tzv. díbitu (dvojici bitů). Podobně u osmistavových modulací přenáší signálový prvek třibit (tři bity) atd.

Přenosová rychlost v_p je definována vztahem pomocí modulační rychlosti:

$$v_p = v_m \cdot \log_2 m \text{ [bit/s]},$$

kde m je počet stavů číselného signálu.

Tak např. pro již zmiňovanou čtyřstavovou modulaci ($m=4$), reprezentuje každý signálový prvek dva bity podle obrázku:



A pro přenosovou rychlost tak platí:

$$v_p = v_m \cdot 2,$$

to znamená, že přenosová rychlost je tedy dvakrát větší než modulační nebo chceme-li, modulační rychlost je poloviční než rychlost přenosová. Zřejmě pro dvoustavovou modulaci bude platit, že modulační rychlost je rovna rychlosti přenosové. Pro přehlednost uvádím vztah obou rychlostí v závislosti na použité modulaci:

$v_m = v_p$	dvoustavová modulace (m=2)
$v_m = v_p/2$	čtyřstavová modulace (m=4)
$v_m = v_p/3$	osmistavová modulace (m=8)
$v_m = v_p/4$	šestnáctistavová modulace (m=16)
atd.	

Z výše uvedeného je jasné, že s vícestavovou modulací nám roste přenosová rychlost (přenášíme více bitů v jednom symbolu) a zároveň snižujeme modulační rychlost, což je výhodné z hlediska možnosti použití užší šířky pásma a tím tak šetříme frekvenční pásmo. Nevýhodou však je, že s rostoucím počtem stavů narůstá při stejném vysílacím výkonu i chybovost v důsledku horšího rozlišení jednotlivých stavů, ale to už je jiná kapitola.

Pro toho, kdo se chce dovědět víc:

Modulační rychlost, měřená v Baudech, stále ještě neposkytuje věrohodný obrázek o míře schopnosti určité přenosové cesty přenášet data. Jestliže na určité konkrétní přenosové cestě lze dosáhnout modulační rychlosti například 600 Baudů, kolik bitů lze přes ni přenést třeba za jednu sekundu? Zapamatujme si dobře, že na takto položenou otázku nelze odpovědět!!!

Záleží totiž na tom, kolik různých alternativ reprezentuje jedna změna skutečně přenášeného signálu. Neboli: když v rámci modulace měníme (skokem) některý z parametrů přenášeného signálu, z kolika možných hodnot volíme? Pokud jen ze dvou, pak každá z nich může reprezentovat jednu binární číslici (jednu 0 či 1), a pak platí „co změna to jeden bit“. Pokud ale volíme ze čtyř možností, pak každá může reprezentovat dvojici binárních číslic, a pak by platilo „co změna, to dva bity“, a stejně tak bychom mohli pokračovat dále. Obecně pak platí, že počet bitů reprezentovaných („nesených“) jednou změnou přenášeného signálu je dvojkový logaritmus počtu možných stavů (možností).

Věrohodnější obrázek o schopnosti konkrétní přenosové cesty přenášet data poskytuje až tzv. přenosová rychlost, měřená v bitech za sekundu. Předchozí úvaha nás opravňuje formulovat následující vzoreček, vyjadřující konkrétní závislost mezi modulační rychlostí a přenosovou rychlostí:

$$v_p = v_m * \log_2 m$$

kde m je zmíněný počet možných stavů přenášeného signálu.

Znovu se vraťme k definicím modulační a přenosové rychlosti. Tzv. modulační rychlost, měřená v Baudech a vyjadřující počet změn přenášeného signálu za jednotku času, nám ještě neříká nic o tom, kolik bitů je tímto signálem přenášeno. Naproti tomu přenosová rychlost, měřená v bitech za sekundu, nám říká kolik datových bitů je možné přenést za časovou jednotku, a naopak nevypovídá nic o způsobu, jakým se toho dosahuje, neboli o četnosti změn přenášeného signálu. Mezi oběma veličinami přitom zdaleka nemusí platit rovnost - pokud se na vyjádření jednoho datového bitu „spotřebují“ dvě změny přenášeného signálu (což je dáno použitým způsobem kódování jednotlivých bitů), pak přenosová rychlost vychází číselně poloviční oproti rychlosti modulační. Naopak, pokud jedna změna přenášeného signálu je změnou mezi čtyřmi možnými stavy, pak každá takováto změna může reprezentovat hned dva datové bity, a přenosová rychlost bude tudíž dvojnásobná oproti rychlosti modulační. Obě rychlosti se přitom číselně rovnají pouze v případě, kdy je přenášený signál pouze dvoustavový, a každá jeho změna tak reprezentuje jeden jediný datový bit.

Rozdíl mezi modulační a přenosovou rychlostí je nejmarkantnější u dnešních telefonních modemů. Ty totiž používají často dosti složité a komplikované metody modulace a kódování, díky tomu dokáží pracovat s relativně velkým počtem možných stavů přenášeného signálu, a svých přenosových rychlostí tudíž dosahují i při relativně nízkých modulačních rychlostech. Na druhou stranu jim jiné řešení nezbyvá, protože mají k dispozici přenosový kanál s pevně danou šířkou přenosového pásma (300 až 3400 Hz, neboli 3,1 kHz), a podle Nyquistova kritéria na něm nemá smysl používat vyšší modulační rychlost než 6200 Baudů (tj. číselně dvojnásobnou oproti šířce pásma). V praxi jsou ale stejně používány ještě nižší modulační rychlosti, jak ukazuje tabulka:

Přenosová rychlost, v bitech za sekundu	Modulační rychlost, v Baudech	m - počet rozlišovaných stavů přenášeného signálu	$\log_2 m$ - počet bitů, reprezentovaných jednou změnou přenášeného signálu	Označení přenosového standardu
2400	600	16	4	V.22bis
9600	2400	16	4	V.32
14400	2400	64	6	V.32bis
28800	2400-3200	512	9	V.34

Z ní vidíme, že například modem s přenosovou rychlostí 14,4 kilobitů za sekundu pracuje se signálem, který se mění 2400-krát za sekundu (tj. má modulační rychlost 2400 Bd), přičemž tento signál může nabývat celkem $n=2^6$ neboli 64 různých hodnot, a jedna změna tohoto signálu tudíž reprezentuje šest bitů.

A ještě jednu poznámku: kdyby vám někdo nabízel modem s přenosovou rychlostí například 2400 Baudů, obraťte se k němu zády. Pleťte si totiž hrušky a jablka.

Informační rychlost

To, co by vám mělo stačit ke státnicím:

Podle pana Kováře je informační rychlost definována jako střední množství informace generované zdrojem zpráv za jednotku času podle vztahu:

$$r(Z) = \lim_{N \rightarrow \infty} \frac{\text{informace nesená zprávou délky } N}{\text{doba trvání zprávy délky } N} = \frac{H(Z)}{T} \quad [\text{bit/s}],$$

kde $H(Z)$ je entropie zdroje zpráv (viz níže). Dále platí vztah, který už nabývá více na svém významu a objasňuje vlastně co to informační rychlost je:

$$v_p \geq r$$

Tento vztah vyjadřuje, že informační rychlost je menší nebo maximálně rovna rychlosti přenosové (rovnost nastává jen v případě nezávislosti symbolů se stejně pravděpodobnými znaky). Jde tedy o to, že při přenosu užitečných (informačních) dat se přenáší i data „pomocná“, a proto je informační rychlost většinou menší než přenosová rychlost. Tak například pokud vám nějaký poskytovatel internetového připojení bude cpát do hlavy možnost připojení rychlostí třeba 1 Mbit/s (= *přenosová rychlost*), zcela

určitě nikdy takovou rychlostí stahovat nebudete. Skutečná nižší rychlost stahování je právě onou *informační rychlostí* (to jsou ty užitečná data, které vás jako uživatele zajímají).

Dalším možným způsobem jak si informační rychlost definovat, je pomocí *rychlosti kódování*, což je parametr charakterizující míru zabezpečení dat při procesu kanálového kódování. Necht' vstupní data jsou členěna do k-bitových skupin (*zdrojová slova*). Toto zdrojové slovo je zakódováno kanálovým kóděrem do *kódového slova* o délce n (platí, že $n > k$). Poměr k/n se označuje jako rychlost kódování a uvádí zároveň i informační rychlost.

Například kód o rychlosti $3/4$ (k/n) odpovídá třem užitečným bitům, které jsou chráněny jedním pomocným bitem.

Pro toho, kdo se chce dovědět víc:

Přenosová rychlost, kterou jsme se zabývali, je veličinou, která říká velmi mnoho o schopnostech konkrétní přenosové cesty přenášet data. V praxi je ovšem nutné mít na paměti, že jde o veličinu nominálního charakteru, která ještě nemusí vypovídat příliš přesně o tom, jaké objemy dokáže příslušná přenosová cesta přenést za delší časové období.

Jeden otrlý počítačový vtíp, pocházející ještě z dob sálové prehistorie, říká že není důležité, kolik instrukcí vykoná určitý počítač resp. procesor za sekundu, ale že skutečně důležité je to, kolik jich udělá za týden, měsíc či rok. Pointa spočívá v tom, že při provozu počítače může docházet k všelijakým výpadkům, odstávkám a dalším vlivům, kvůli kterým procesor buď vůbec nevykonává žádné instrukce (nebo je i vykonává, ale tyto jdou na vrub různým režijním činnostem, jako třeba diagnostice). Celý vtíp pak názorně dokumentuje rozdíl mezi veličinou nominálního charakteru, jakou je třeba zmíněný počet instrukcí vykonaných za sekundu, a veličinou „efektivní“, která vyjadřuje skutečný (užitečný) efekt, obvykle sledovaný za určité delší období.

Počet instrukcí, vykonaných procesorem za jednu sekundu, je třeba chápat především jako údaj vyjadřující jak dlouho trvá provedení jedné instrukce - v tom smyslu, že když procesor provádí své strojové instrukce rychlostí 1000 instrukcí za sekundu (například), pak provedení jedné strojové instrukce trvá jednu tisícinu sekundy. Ani v námi zvoleném případě procesoru a jeho instrukcí pak není možné použít jednoduchou násobilku a říct, že za sekundu jich provede přesně tisíc. Do hry totiž vstupují různé vlivy, jako například DMA přenosy či výpadky bloků ve vyrovnávacích pamětech, a ty mohou významněji prodlužovat dobu kterou trvá provedení jednotlivých instrukcí. Pokud bychom tedy sledovali počet skutečně provedených instrukcí za delší časové období, počítali je a pak vztáhli na časovou jednotku (například sekundu), dostali bychom nejspíše poněkud jiné číslo. Hlavně bychom ale dostali veličinu „efektivního“ (průměrného, sumárního) typu, která již bere do úvahy i další faktory než jen nominální rychlost provádění instrukcí.

Srovnáme-li si obě veličiny - nominální rychlost provádění instrukcí a efektivní počet skutečně provedených instrukcí - pak rozměrově budou shodné, neboť obě vyjadřují počet provedených instrukcí za sekundu. Číselně se ale mohou i dosti významně lišit, a tato jejich odlišnost bude odrazem nejrůznějších vlivů, které se v praxi uplatňují. Je velmi důležité si uvědomit, že tyto vlivy mohou působit oběma směry - mohou nejen číselně snižovat „efektivní“ rychlost oproti rychlosti nominální, ale mohou ji stejně tak i zvyšovat. Námi zvolený příklad s prováděním strojových instrukcí to umožňuje názorně dokumentovat: dnešní procesory velmi často používají techniky tzv. prokládání, a zpracovávají více instrukcí najednou - například jednu právě začínají, druhou mají rozpracovanou zhruba z poloviny a třetí právě dokončují. Jestliže takto dochází k trojnásobnému prokládání, tj. procesor v každém okamžiku pracuje na třech instrukcích současně a provedení každé instrukce trvá jednu tisícinu sekundy, pak za tuto jednu tisícinu sekundy procesor stihne dokončit hned tři instrukce, a za jednu sekundu tak teoreticky stihne provést tři tisíce instrukcí (odmyslíme-li si na chvíli efekt opačně působících vlivů). Takže „efektivní“ rychlost může být dokonce i výrazně vyšší než rychlost nominální.

Přenosová rychlost vs. informační rychlost

Podobně jako k rychlosti provádění strojových instrukcí musíme při datových přenosech přistupovat i k rychlosti přenosové. Jde opět o veličinu nominálního charakteru, kterou je vhodné chápat jako vyjádření toho, jak dlouho trvá přenos jedné elementární jednotky informace, tedy jednoho bitu. Takže když se například někde něco přenáší rychlostí 10 Mbps (megabitů za sekundu), pak je třeba tomu rozumět tak, že přenos jednoho bitu zabere jednu desetimilióntinu sekundy. Nikoli tak, že za jednu sekundu se přenesou 10 „mega“ bitů, za deset sekund 100 „mega“ atd. Mějme to na paměti, například i v souvislosti s Ethernetem. „Efektivní“ alternativou k nominální přenosové rychlosti je v datových přenosech veličina, označovaná nejčastěji jako „přenosový výkon“ (anglicky: throughput, doslova: propustnost), ale někdy také jako „**informační rychlost**“. Tento druhý možný název šikovně vystihuje jednu důležitou specifiku datových přenosů: aby bylo možné přenést určitý počet „užitečných“ bitů (resp. bitů vyjadřujících určitou informaci), je nutné ve skutečnosti přenést navíc i určité režijní bity, nutné pro korektní fungování přenosových mechanismů. Jde například o nejrůznější kontrolní součty, sloužící potřebám detekce neporušenosti datového obsahu při spolehlivých přenosech, či o různé adresy příjemců a odesílatelů, identifikátory obsahu, hodnoty čítačů vyjadřujících pořadí atd.

Přenosová rychlost existenci těchto režijních bitů nebere v úvahu, zatímco informační rychlost ano. Přenosová rychlost říká, jak dlouho trvá přenos jednoho bitu a nezajímá se o to, zda jde o režijní bit či o bit nesoucí „užitečnou“ informaci. Naproti tomu informační rychlost obě situace již rozlišuje, všímá si pouze „užitečných bitů“ nesoucích informace, a vyjadřuje tudíž množství „užitečných“ dat, které je daná přenosová cesta schopna skutečně přenést. Navíc jde o veličinu „efektivního“ charakteru, která bere v úvahu průměrné chování vztahované k delším časovým úsekům.

Problém je ovšem v tom, že není nijak lehké exaktně definovat, co jsou „užitečná“ data, a co jsou data režijní. Proto je obtížné přesně definovat i samotnou informační rychlost, i když při intuitivním pohledu je to vcelku jednoduché. To je zřejmě také hlavní důvod, proč se v běžné praxi s informační rychlostí moc npracuje, a místo toho se používá především přenosová rychlost. Může to ale mít ostatně i trochu jiné důvody než jen obtížnost a vážnost definice informační rychlosti: říci například o Ethernetu, že pracuje s přenosovou rychlostí 10 Mbps, je jistě působivější než prozradit, že jeho informační rychlost v reálných podmínkách bývá někde mezi 3 až 4 Mbps.

Faktory ovlivňující informační rychlost

Jak jsme si již naznačili, na informační rychlost mají vliv různé „režijní“ bity, přenášené spolu s bity „užitečnými“. Kromě toho je zde ale celá řada dalších vlivů, které působí směrem ke snižování informační rychlosti vůči nominální přenosové rychlosti. Jde například o asynchronní způsob komunikace, při kterém mohou mezi jednotlivě přenášenými částmi dat (jednotlivými znaky, bloky atd.) existovat libovolně dlouhé prodlevy. Ty nominální přenosová rychlost nebere do úvahy, zatímco informační rychlost ano.

Informační rychlost dále bere do úvahy například fungování mechanismů zabezpečujících spolehlivost přenosu - dojde-li například k poškození přenášeného bloku a ten musí být přenesen znovu, na přenosovou rychlost to nemá vliv, ale informační rychlost to samozřejmě zřetelně zmenšuje.

Další zajímavý vliv je možné dokumentovat na již citovaném příkladu Ethernetu. Ten počítá s existencí tzv. sdíleného přenosového média, do kterého může v principu vysílat kdokoli, ale v praxi musí zaučinkovat vhodný mechanismus, který v případě vícenásobného zájmu rozhodne, kdo bude moci vysílat a kdo musí počkat. Ethernet používá pro tyto účely přístupovou metodu CSMA/CD, jejíž samotná podstata také zavádí určitou režii snižující informační rychlost oproti nominální přenosové rychlosti. Navíc jde o tzv. neřízenou přístupovou metodu, a její režie při zvyšující se zátěži významně vzrůstá.

Možnosti zvyšování informační rychlosti

Vedle celé řady „režijních“ faktorů, která v datových komunikacích působí směrem ke snižování informační rychlosti oproti přenosové rychlosti, však existují i takové vlivy, které působí opačně. Jde

zejména o mechanismy on-line komprese, zabudovávané do dnešních telefonních modemů. Například modemy dle standardu MNP 4 dokáží komprimovat data v poměru 2:1, a modemy využívající standard V.42bis dokonce v poměru 4:1. Jednoduchým výpočtem pak lze zjistit, že například pro modem o přenosové rychlosti 28,8 kbps by při maximální kompresi mohl teoreticky dosahovat efektivní informační rychlosti 115,2 kbps (ovšem bez uvážení vlivu opačně působících faktorů).

V praxi je ovšem situace poněkud střídavější - komprimační mechanismy, zabudované do modemů, totiž musí fungovat v reálném čase, a nemohou tudíž být příliš efektivní. Také citovaná úroveň komprese (až 4:1) je spíše horním odhadem, zatímco reálně dosahované úrovně komprimace bývají nižší. Záleží přitom i na konkrétní povaze dat, která jsou přenášena - jde-li například o soubor, který byl zkomprimován již před svým přenosem (některým výkonnějším kompresním prostředkem), pak pokus modemu o jeho další kompresi pomocí „hloupějších“ kompresních prostředků vede ve skutečnosti na jeho zvětšení, a tudíž na faktické snížení informační rychlosti.

Shannon – Hartleyův zákon

To, co by vám mělo stačit ke státnicím:

Tento zákon nám říká, jakou maximální přenosovou kapacitu C můžeme přenést v určitém rádiovém komunikačním systému. Libovolný rádiový komunikační systém nemůže přenést v určitém časovém intervalu zcela libovolné, neomezené množství informace, nýbrž pouze množství nepřesahující jeho *přenosovou kapacitu* C . V každém reálném systému je totiž přítomen šum a řada dalších rušivých jevů, které ztěžují na přijímací straně vyhodnocování relativně velmi malých užitečných signálů.

Předpokládejme dále, že v radiokomunikačním kanálu působí pouze aditivní bílý Gaussovský šum AWGN. Přenosová kapacita C je potom definována jako maximální množství informace vyjádřené v bitech, jež může být daným systémem přeneseno za určitý časový interval (nejčastěji za 1 sekundu), a to při nulové (přesněji řečeno libovolně malé) chybovosti BER (*pozn.:* pokud se snažíme tuto maximální přenosovou kapacitu C překročit, chybovost prudce narůstá).

Označíme-li si střední hodnotu výkonu užitečného signálu na vstupu přijímače symbolem S a střední hodnotu výkonu šumu symbolem N , dále šířku pásma daného kanálu B , je pak přenosová kapacita C určena *Shannon-Hartleyovým vztahem*:

$$C = B \log_2 \left(1 + \frac{S}{N} \right), \text{ resp. } C = 3,32B \log_{10} \left(1 + \frac{S}{N} \right) \text{ [bit/s]}$$

Přenosová kapacita C vyjadřuje maximální dosažitelnou rychlost přenosu informace **idealizovaným** radiokomunikačním systémem, v němž působí jen výše zmíněný aditivní bílý Gaussovský šum AWGN, a to při chybovosti BER blížící se nule. Dále se předpokládá, že je použito *optimální kódování* a *optimální modulace*.

Reálné systémy se tedy mohou této kapacitě pouze přiblížit, a to tím dokonaleji, čím věrněji se v nich použité metody kódování a modulace přibližují metodám optimálním. Přitom je nutné zdůraznit, že tyto dokonalé optimální metody nejsou v praxi dosažitelné, přičemž pokusy o jejich realizaci vedou k rapidnímu nárůstu složitosti příslušných technických prostředků.

Z výše uvedeného vztahu je jasné, že požadovanou kapacitu C je možné dosáhnout různými kombinacemi parametrů B , S a N . U pozemských radiokomunikačních systémů ji lze často získat použitím velkých výkonů vysílačů, antén s velkým ziskem apod., tedy velkým poměrem S/N , takže se zde vystačí s relativně malými šířkami pásma B . Naproti tomu například pro družicové systémy jsou typické malé výkony palubních vysílačů, takže požadovanou kapacitu C je možné získat jedině náležitým rozšířením pásma B .

Možná je ještě dobré si trochu poupravit Shannon-Hartleyův vztah a ukázat, že s rostoucí šířkou pásma B nám pro stejnou hodnotu přenosové kapacity C postačuje menší poměr S/N :

$$C = B \log_2 \left(1 + \frac{S}{N}\right) \Rightarrow C \approx \frac{B}{\ln(2)} \frac{S}{N} \Rightarrow \frac{S}{N} \approx \frac{C \cdot \ln(2)}{B}$$

Pro toho, kdo se chce dovědět víc:

Zamysleme se nyní nad jednou velmi důležitou otázkou, která je zvláště aktuální u komutovaných (vytáčených) linek veřejné telefonní sítě: máme-li k dispozici určitou přenosovou cestu s jejími konkrétními a neměnnými obvodovými vlastnostmi, můžeme na ní dosáhnout libovolně vysoké přenosové rychlosti? Neboli, řečeno jinými slovy: budeme-li zdokonalovat technickou stránku přenosu a díky tomu zvyšovat přenosovou rychlost na určité přenosové cestě, budeme to moci dělat libovolně dlouho, nebo někde narazíme na nějakou z principu nepřekonatelnou bariéru? Nebo ještě jinak: když dnes existují modemy pro komutované linky veřejné telefonní sítě pracující s přenosovou rychlostí 28,8 kilobitů za sekundu, má smysl ještě chvíli počkat, až se na trhu objeví třeba modemy s rychlostí 64 kbps, 128 kbps apod.?

Odpověď začneme hledat vzorečkách a vztazích, které jsme si již dříve naznačili: jestliže maximální modulační rychlost je podle Nyquistova kritéria dvojnásobná oproti dostupné šířce přenosového pásma, a tato šířka přenosového pásma je pro danou přenosovou cestu fixována (je neměnná), pak z toho jednoduše vyplývá, že modulační rychlost nelze libovolně dlouho zvyšovat (a její maximální hodnota je také pevně dána). Jestliže přenosová rychlost závisí na modulační rychlosti podle již uvedeného vzorečku:

$$v_p = v_m \cdot \log_2 m$$

pak poslední možností pro zvyšování přenosové rychlosti je zvyšování parametru m , neboli zvyšování počtu rozlišovaných stavů přenášeného signálu. Výsledná přenosová rychlost by při lineárním zvyšování m sice rostla pomaleji (logaritmicky), ale přesto bychom se při dostatečně vysoké hodnotě m mohli dostat s přenosovou rychlostí tak vysoko, jak potřebujeme.

Podívejme se ale na tuto možnost nejprve obyčejným „selským rozumem“: budeme-li zvyšovat počet možných stavů přenášeného signálu, bude čím dál tím těžší je správně rozpoznat, resp. rozlišit od sebe. Intuitivně je tedy vcelku zřejmé, že něco takového nemůžeme dělat libovolně dlouho, ale že dříve či později narazíme na mez, za kterou už příjemce nebude schopen dostatečně přesně rozlišit stavy přijímaného signálu.

Zajímavou otázkou ovšem je, zda tato mez je dána našimi momentálními schopnostmi, resp. dokonalostí přenosové techniky a je možné očekávat její postupné posouvání, nebo zda jde o mez závislou na něčem jiném, co nemá s dokonalostí dostupné techniky nic společného (a co se tudíž nemusí posunout ani při sebedokonalejší technice).

Odpověď je (bohužel) taková, že zmíněná hranice je principiálního charakteru, a je nezávislá na dokonalosti naší techniky a technologie. Jinými slovy: i kdyby se výrobci modemů snažili sebevíce, přes onu magickou hranici se nikdy nedostanou.

Ale kde ona hranice leží? Tím, kdo tuto hranici našel (v roce 1948), byl zakladatel moderní teorie informace, pan Claude Shannon. Ten totiž zjistil, že maximální dosažitelná přenosová rychlost závisí jednak na dostupné šířce přenosového pásma (což je ihned zřejmé), ale pak už jen na „kvalitě“ přenášeného signálu, vyjádřené tím jak dobře jej lze odlišit od nepříznivých vlivů, zejména všudypřítomného šumu. Konkrétní vzoreček závislosti přenosové rychlosti na uvedených veličinách, označovaný také jako tzv. Shannonův teorém, je následující:

$$\text{maximální } C = \text{šířka pásma} \cdot \log_2(1 + \text{signál/šum})$$

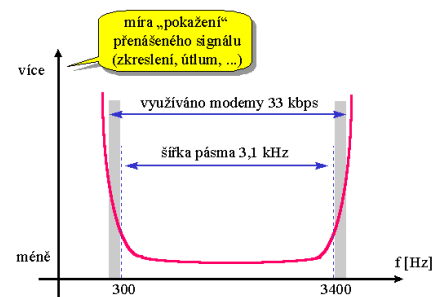
Přitom poměr „signál/šum“ (též: odstup signálu od šumu) je veličina, která je opět dána reálnými obvodovými vlastnostmi konkrétní přenosové cesty, a v praxi většinou není možné ji výrazněji ovlivnit (vyjadřuje totiž míru toho, jak se do „užitečného“ signálu přimíchávají jiné signály rušivého charakteru). Například kvalitní komutovaná linka analogové veřejné telefonní sítě dosahuje odstupů signál/šum 1000:1. Dosazením této hodnoty do Shannonova vzorečku (spolu s šířkou přenosového pásma 3,1 kHz) nám vyjde, že maximální dosažitelná přenosová rychlost na běžných komutovaných linkách veřejné telefonní sítě je kolem 30 000 bitů za sekundu!!

Kupte si perpetuum mobile!

Uvědomme si dobře, co právě vyslovené tvrzení znamená: Shannonův teorém je zcela nezávislý na technické dokonalosti - nenajdete v něm ani vliv použité modulace, ani vliv použitého kódování. To ale znamená, že sebedokonalější technika přenosu dat nemůže při pevně dané šířce pásma a kvalitě přenosu (dané odstupem signálu od šumu) překročit mez danou Shannonovým teorémem. Snahy překonat tuto mez pak mají stejnou šanci na úspěch, jako snahy sestrojít perpetuum mobile.

Jaká je ale současná praxe v oblasti telefonních modemů pro komutované linky veřejné telefonní sítě? Dnes již existují (a jsou běžně k dostání) modemy pro přenosovou rychlost 28,8 kbps, které se teoretické hranici vyplývající z Shannonova teorému velmi blíží. Svědčí to mimo jiné i o vyspělosti naší současné techniky.

Na trhu však již dnes jsou i modemy, dosahující přenosové rychlosti 33 kilobitů, což je nad hranici danou Shannonovým teorémem (a nejde přitom o podvod). Podařilo se tedy zkonstruovat perpetuum mobile, resp. vyvrátit Shannonův teorém? Nikoli, to skutečně nejde. Zmíněné modemy pro 33 kbps plně respektují Shannonův teorém, neboť používají o něco větší šířku pásma než původních 3,1 kHz. Dokáží totiž využít i okrajové části přenosového spektra běžných komutovaných linek veřejné telefonní sítě (viz obrázek), které již mají natolik špatné přenosové vlastnosti, že pro ostatní modemy nejsou použitelné - vlastně si tím dokáží „roztáhnout“ původní přenosové pásmo o šířce 3,1 kHz.



Propustnost kanálů

To, co by vám mělo stačit ke státnicím:

Dnes existuje spousta různých typů signálů, které mají různé parametry a charakteristiky. Ať už máme ale jakýkoliv signál, musíme být schopni ho přenést telekomunikačním kanálem. Abychom mohli zkoumat propustnost jednotlivých kanálů, musíme se nejprve věnovat signálům. Objem signálu je charakterizován třemi základními prvky:

- **Dynamický rozsah signálu** – např. u hovorového signálu rozlišujeme dvě mezní úrovně, a to *šepot* a *nejhlasitější* složku hovorového signálu (výkřik), z praktického hlediska definujeme poměr střední hodnoty výkonu signálu ku střední hodnotě výkonu šumu
- **Šířka pásma signálu** – např. u akustického signálu nás zajímá *nejnižší* a *nejvyšší* frekvence, které je lidské ucho schopno vnímat
- **Doba trvání signálového prvku** – je to nejmenší část signálu, která je samostatně rozlišitelná (např. bit)

Aby byl tedy signál přenesen bez nějakého zkreslení či dokonce ztráty informace, musí být **propustnost kanálu vždy větší nebo maximálně rovna objemu signálu**.

Pozn.: Počet typů signálů je mnohem větší než počet typů kanálů.

Rozlišujeme primárně tři *typy kanálů*, které propouštějí různé signály, z toho také plynou různé šířky pásma, které používají:

- Telefonní kanál (0,3 ÷ 3,4 kHz)
- Rozhlasový kanál (40 Hz ÷ 15 kHz)
- Televizní kanál (50 Hz ÷ 6 MHz)

Entropie

To, co by vám mělo stačit ke státnicím:

Informaci, kterou chceme přenést od vysílače k přijímači, vyjadřujeme ve formě *zprávy* (ta je vytvořena zdrojem zpráv). Zpráva se skládá z jednotlivých *prvků* zprávy (např. zpráva ve formě nějakého čísla je vyjádřena pomocí deseti číslic 0-9). Zpráva má z hlediska příjemce náhodný charakter. Zpráva o velmi pravděpodobném jevu nese malé množství informace, naopak zpráva o málo pravděpodobném jevu nese velké množství informace.

Jakýsi průměr, tedy **průměrné množství informace**, které je zprávou nesené, nazýváme právě **entropie**. Jednotkou entropie je *Shannon* [Sh].

Obecně je entropie definována následovně:

$$H = -\sum_{i=1}^M p_i \cdot \log_2 p_i ,$$

kde p_i je pravděpodobnost vyslání prvku i . Tento vztah platí, mají-li jednotlivé prvky různou pravděpodobnost výskytu. Pokud mají všechny prvky ve zprávě stejnou pravděpodobnost výskytu, pak je entropie definována jednodušeji:

$$H = \log_2 M ,$$

kde M představuje počet prvků abecedy, z nichž každý má stejnou pravděpodobnost.