



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

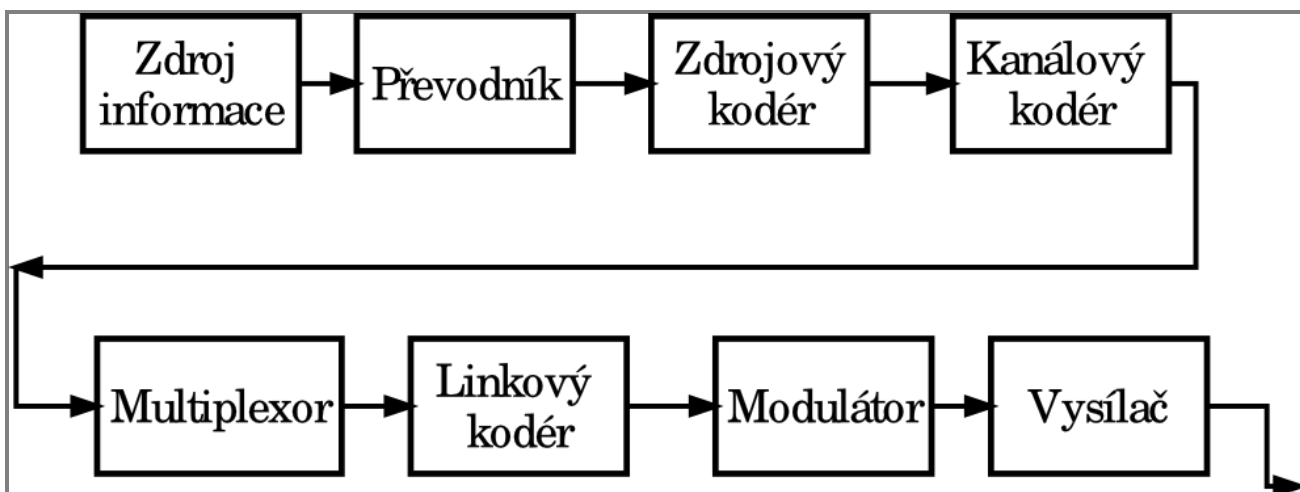
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

PSK2-1

Název školy:	Vyšší odborná škola a Střední průmyslová škola, Božetěchova 3
Autor:	Ing. Marek Nožka
Anotace:	Zdrojové kódování a komprese dat
Vzdělávací oblast:	Informační a komunikační technologie
Předmět:	Počítačové sítě a komunikační technika (PSK)
Tematická oblast:	Vrstvy protokolu TCP/IP
Výsledky vzdělávání:	Žák objasňuje princip komprese dat
Klíčová slova:	zdrojové kódování, komprese dat, Huffmanovo kódování
Druh učebního materiálu:	Online vzdělávací materiál
Typ vzdělávání:	Střední vzdělávání, 3. ročník, technické lyceum
Ověřeno:	VOŠ a SPŠE Olomouc; Třída: 3L
Zdroj:	Vlastní poznámky, Wikipedia, Wikimedia Commons

Zdrojové kódování

Digitální přenosový řetězec



Blokové schéma vysílací strany přenosového řetězce

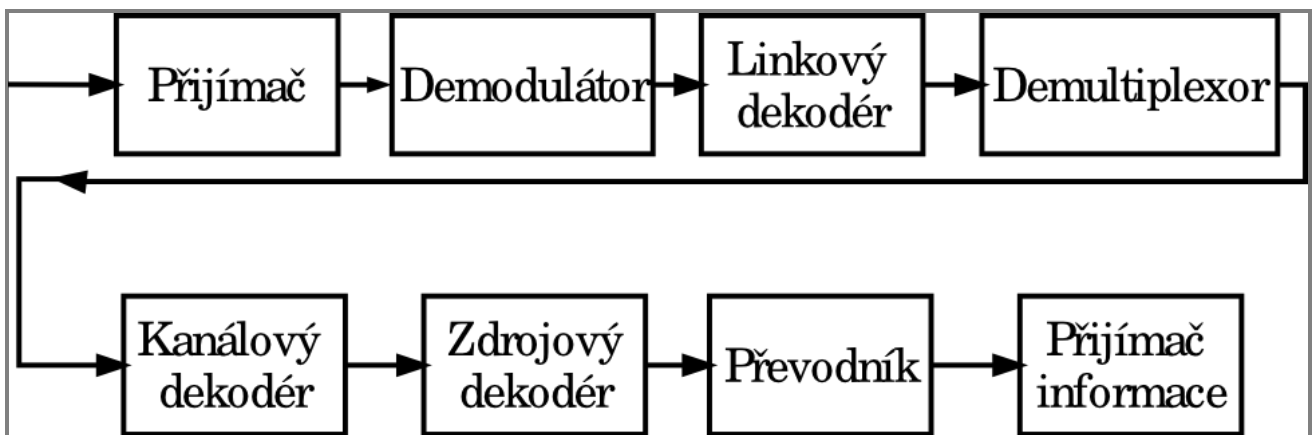
Blokové schéma vysílací strany přenosového řetězce je na obrázku.

Zdroj informace dodává signál. Informacemi, které můžeme přenášet, mohou být: lidský hlas, hudba, údaj o teplotě vzduchu, atd. atd.

Převodník potom signál ze zdroje informace převádí na potřebný typ binárního signálu. Převodníkem může být mikrofon, teploměr, atd, s A/D převodníkem.

Tento binární signál, je po dalších úpravách ve vysílači zpracováván tak, aby mohl být co možná *nejhospodárněji* přenesen k přijímací straně. Ta se skládá z obdobných bloků jako strana vysílací, ale jejich funkce je přesně opačná.

Mezi vysílačem a přijímačem je konkrétní přenosové médium. Například světlo v optickém vlákně, nebo elektromagnetická vlna, která se šíří volným prostorem.



Blokové schéma přijímací strany přenosového řetězce

Zdrojové kódování slouží ke zvýšení hospodárnosti přenosu. Jeho úkolem je snížit objem přenášených dat na minimum. Jde tedy o **kompresi** -- stlačení, snížení objemu dat. Ta může být *ztrátová* nebo *bezztrátová*.

Bezztrátová komprese

Bezztrátová komprese odstraňuje z dat tzv. redundance (redundantní znamená nadbytečný, hojný)

Při použití bezztrátové komprese lze vždy data přesně rekonstruovat do původní podoby. Nedochozí k žádnému zkreslení.

Rekonstruovaná a původní data jsou stejná.

Redundance

Různé části přenášených dat nesou různou míru informace. Každá zpráva obsahuje znaky, které se více nebo méně často opakují. Čím se daný znak opakuje častěji, tím je běžnější, a tím méně informace nese. Naopak znaky, které jsou ve zprávě vzácné, nesou více informace.

Pokud by byl přidělen všem písmenům stejný objem dat, byl by přenos velice nehospodárný. Logičtější je přidělit písmenům, která nesou více informace, větší objem dat a písmenům, která nesou méně informace, menší objem dat.

Nejznámějším kódem pro zvýšení hospodárnosti přenosu (snížení

redundance) je Morseova abeceda. Ta přiděluje nejkratší značky nejčastěji se vyskytujícím písmenům. Naopak nejdelší značky mají písmena která se v textu objevují nejméně.

. E H
- T - . . - X
.. I - - . . Z
. - A - . - - Y

Huffmanovo kódování

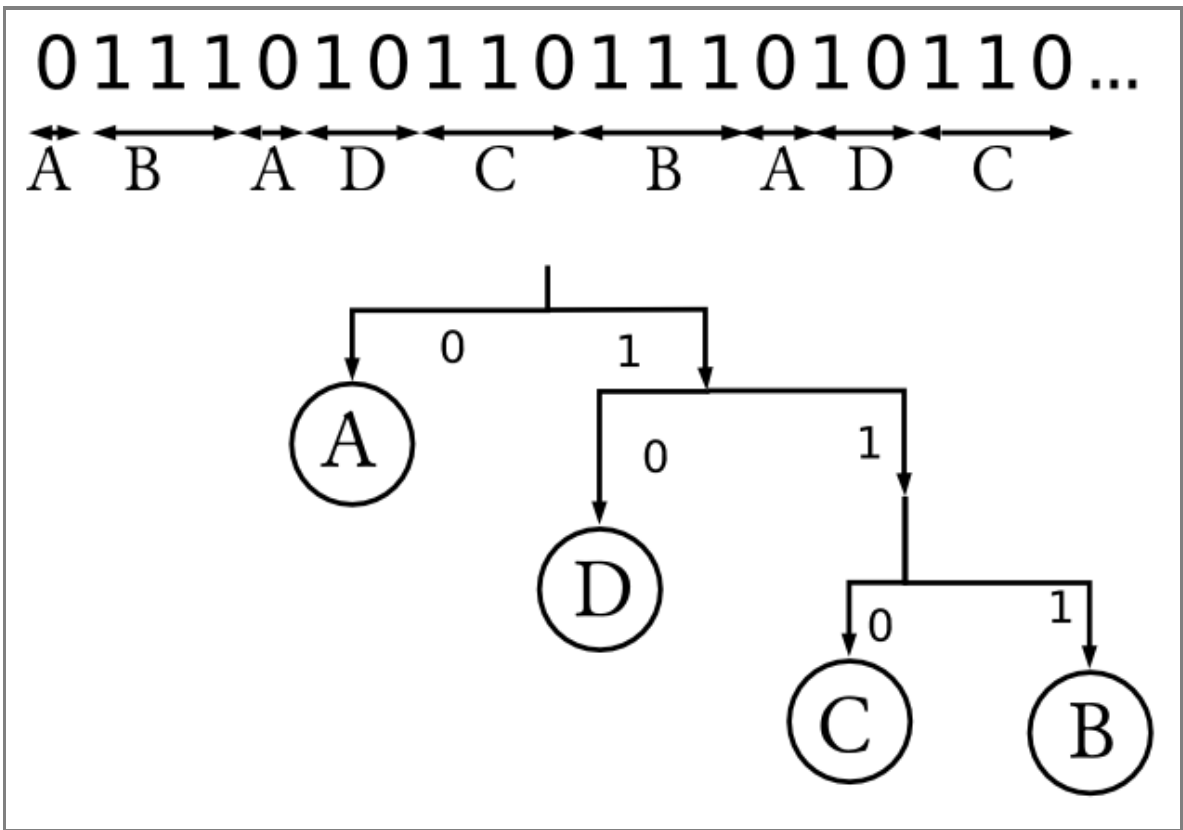
Typickým příkladem kódu pro snížení redundance je Huffmanův kód. Následující příklad chce ukázat, jak se při tvorbě tohoto kódu postupuje.

Mějme abecedu, která obsahuje čtyři znaky: A, B, C a D. Každý z těchto prvků se ve zprávě vyskytuje s jinou četností. To lze vyjádřit pravděpodobností jeho výskytu $P(X)$.

1. prvky seřadíme podle pravděpodobnosti výskytu
2. vybereme dva prvky s nejmenší pravděpodobností, pravděpodobnějším z nich přiřadíme 1, méně pravděpodobnému 0
3. tyto dva prvky s nejmenší pravděpodobností sloučíme do jediného prvku
4. tento postup stále opakujeme
5. značky čteme zprava doleva

A 40% A 40% A 40% BCD 60% 1
B 20% D 25% BC 35% 1 A 40% 0
C 15% B 20% 1 D 25% 0
D 25% C 15% 0

Tímto postupem přidělíme každému znaku jinak dlouhou značku. Vytvoříme tak vlastně strom a při dekódování tímto stromem procházíme. Vždy, když narazíme na nulu, ukončíme čtení jednoho znaku.



Dekódování Huffmanova kódu