Data Communication

BCE 6th Semester

Er. Anuj Sherchan

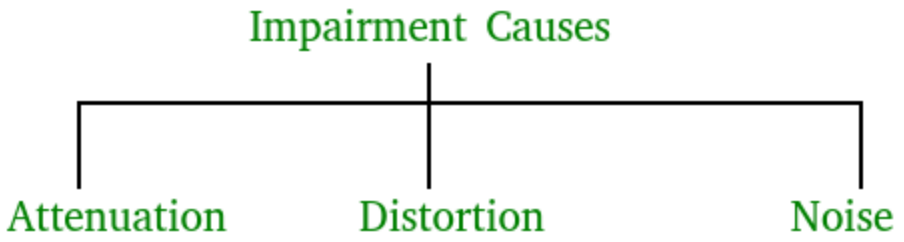
Assistant Professor, Pokhara Engineering College

Unit 7: Impairments, Error Handling and Compression Techniques

- Outline:
- Attenuation, Distortion, Noise, Interference
- Types of Error, its error detection, and correction methods
- Types of Data Compression Techniques

Transmission Impairment

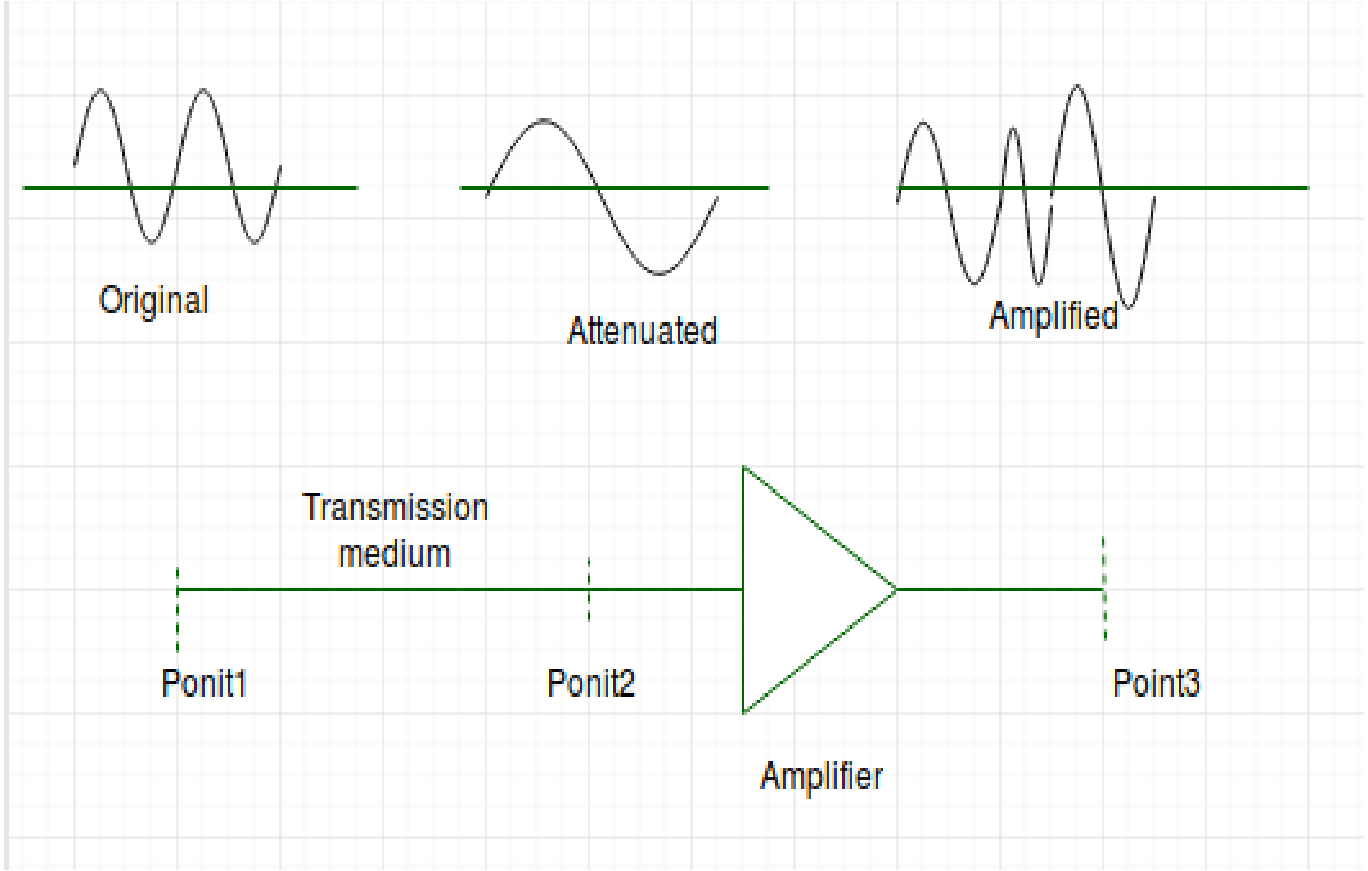
- In a communication system, analog signals travel through transmission media, which tends to deteriorate the quality of the analog signal, which means that the signal at the beginning of the medium is not the same as the signal at the end of the medium.
- The imperfection causes signal impairment.
- Below are the causes of the impairment.



Transmission Impairments

- **Attenuation**
- It means loss of energy.
- The strength of the signal decreases with increasing distance which causes a loss of energy in overcoming the resistance of the medium.
- This is also known as an attenuated signal.
- Amplifiers are used to amplify the attenuated signal which gives the original signal back and compensates for this loss.

Attenuation



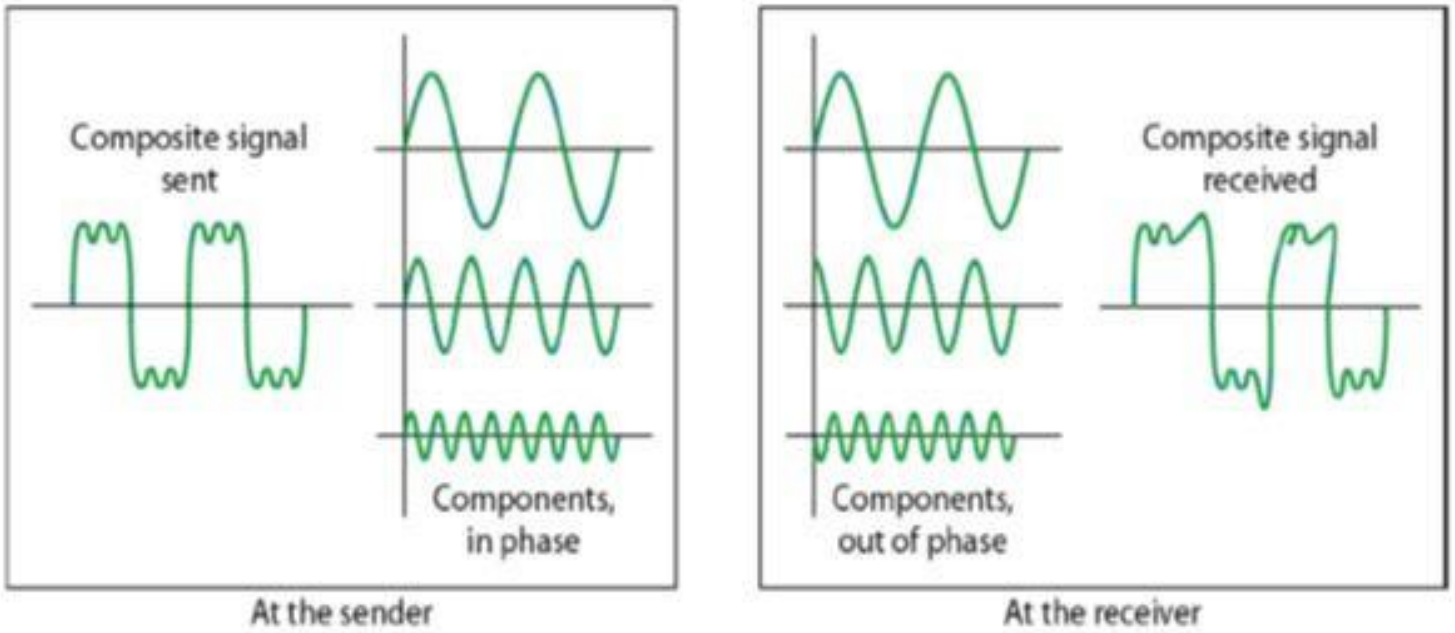
Attenuation

•Attenuation is measured in **decibels(dB)**.
•It measures the relative strengths of two signals or one signal at two different points.
Attenuation(dB) = 10log10(P2/P1)
•P1 is the power at sending end and P2 is the power at receiving end.
• Somewhere the decibel is also defined in terms of voltage instead of power
•In this case because power is proportional to the square of the voltage the formula is
Attenuation(dB) = 20log10(V2/V1)
• V1 is the voltage at sending end and V2 is the voltage at receiving end.

Transmission Impairments

- **Distortion**
- It means changes in the form or shape of the signal.
- This is generally seen in composite signals made up of different frequencies.
- Each frequency component has its own propagation speed traveling through a medium.
- And that is why there is a delay in arriving at the final destination
- Every component arrives at a different time which leads to distortion.
- Therefore, they have different phases at the receiver end from what they had at the sender’s end.

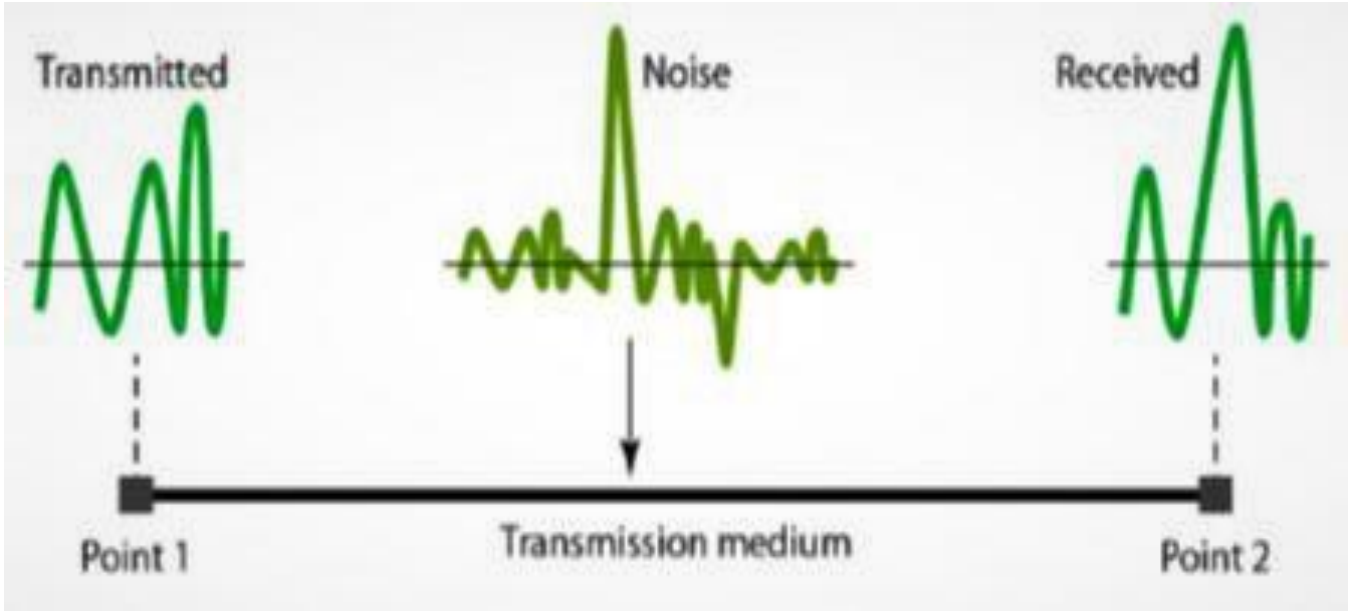
Distortion


Composite signal
sent
Components,
in phase
At the sender
Composite signal
received
Components,
out of phase
At the receiver

Transmission Impairments

- **Noise**
- The random or unwanted signal that mixes up with the original signal is called noise.
- There are several types of noise such as induced noise, crosstalk noise, thermal noise, and impulse noise which may corrupt the signal.
- **Induced** noise comes from sources such as motors and appliances. These devices act as sending antenna and the transmission medium act as receiving antenna.
- **Thermal** noise is the movement of electrons in the wire which creates an extra signal.
- **Crosstalk** noise is when one wire affects the other wire.
- **Impulse** noise is a signal with high energy that comes from lightning or power lines

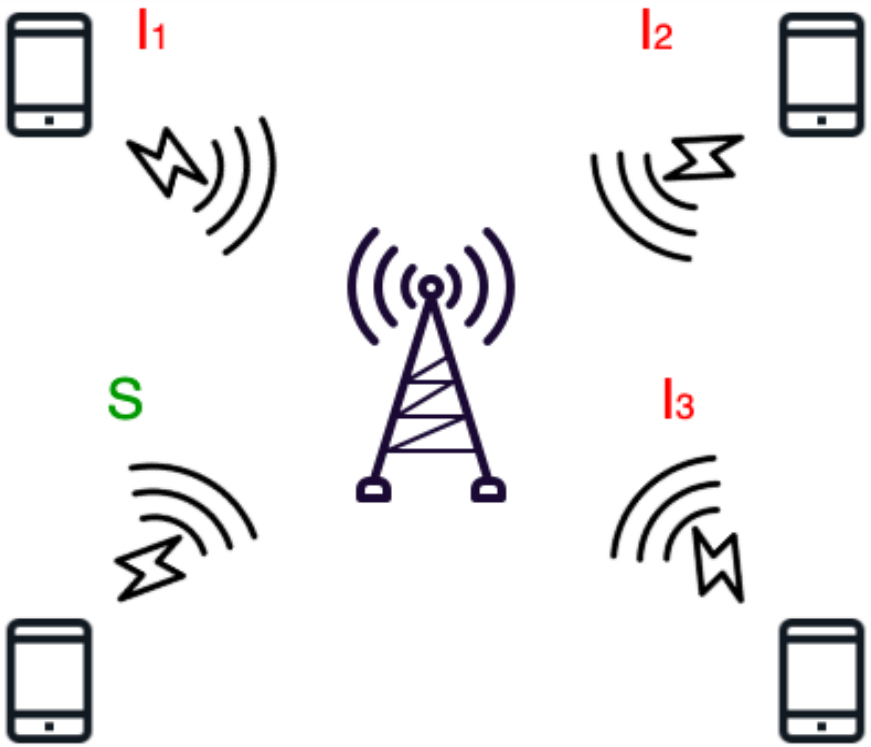
Noise


Transmitted
Noise
Received
Point 1
Transmission medium
Point 2

Transmission Impairments

- In telecommunications, an **interference** is that which modifies a signal in a disruptive manner, as it travels along a communication channel between its source and receiver.
- The term is often used to refer to the addition of unwanted signals to a useful signal.
- Common examples include:
- Electromagnetic interference (EMI)
- Co-channel interference (CCI), also known as crosstalk
- Adjacent-channel interference (ACI)
- Inter symbol interference (ISI)

Interference



The receiver is getting data from the Interferer I1, I2 and I3 and the desired signal S at the same time. The stronger the signals from the interferer is, the worse the SINR or BER becomes.

Transmission Impairments

- **Crosstalk**
- is a disturbance caused by the electric or magnetic fields of one telecommunication signal affecting a signal in an adjacent circuit.
- Essentially, every electrical signal has a varying electromagnetic field.
- Whenever these fields overlap, unwanted signals -- capacitive, conductive, or inductive coupling -- cause electromagnetic interference (EMI) that can create crosstalk.
- **Crosstalk in telephony**
- The definition of crosstalk, as it relates to telecommunication or telephony, is when there is leakage from a separate conversation from a nearby circuit into the phone conversation of someone else nearby.
- The crosstalk issue can be extremely disruptive, particularly in a business setting.
- If it's an analog connection, twisted pair cabling can often be employed to reduce the likelihood of crosstalk.

Transmission Impairments

- **Near-end crosstalk (NEXT)**
- NEXT refers to a cable's ability to reject crosstalk. In other words, the higher the NEXT value, the better the connection's ability to reject crosstalk.
- It is referred to as "near-end" because the interference between the cables is measured at the same end of the cable that is introducing the interference.
- **Power sum near-end crosstalk (PS NEXT)**
- PSNEXT is a NEXT metric that denotes the sum of crosstalk attribution from all adjacent pairs as the sum of the NEXT of the three-wire pairs as they impact the fourth pair in a four-pair cable system.

Transmission Impairments

- **Far-end crosstalk (FEXT)**
- FEXT is the measure of interference between two pairs of a cable. It is determined at the "far end" of a cable with an interfering transmitter.
- **Equal level far-end crosstalk (ELFEXT)**
- ELFEXT is the measure of the FEXT that contains attenuation compensation.
- **Alien crosstalk (AXT)**
- AXT is a measure of interference created by non-related cables routed in close proximity to the cable of interest.

Channel Capacity

- Channel capacity is the maximum information rate that a channel can transmit.
- Two theoretical formulas were developed to calculate the data rate.
- Nyquist for a noiseless channel
- Shannon for a noisy channel

Channel Capacity

- **Noiseless Channel: Nyquist Bit Rate**
- For a noiseless channel, the Nyquist bit rate formula defines the theoretical maximum bit rate.
- ***Nyquist*** proved that if an arbitrary signal has been run through a low-pass filter of bandwidth, the filtered signal can be completely reconstructed by making only 2*Bandwidth (exact) samples per second.
- Sampling the line faster than 2*Bandwidth times per second is pointless because the higher-frequency components that such sampling could recover have already been filtered out.

Nyquist Rate

- If the signal consists of L discrete levels, Nyquist's theorem states:
- Bit Rate = 2 * Bandwidth * $\log_2(L)$ bits/sec
- bandwidth is the bandwidth of the channel,
- L is the number of signal levels used to represent data,
- Bit Rate is the bit rate in bits per second.
- Bandwidth is a fixed quantity, so it cannot be changed.
- Hence, the data rate is directly proportional to the number of signal levels.

Nyquist rate

- **Examples:**
- **Input1:** Consider a noiseless channel with a bandwidth of 3000 Hz transmitting a signal with two signal levels. What can be the maximum bit rate?
  Output1 : Bit Rate = 2 * 3000 * $\log_2(2)$ = 6000bps
- **Input2:** We need to send 265 kbps over a noiseless channel with a bandwidth of 20 kHz. How many signal levels do we need?
  Output2 : 265000 = 2 * 20000 * $\log_2(L)$
  $\log_2(L) = 6.625$
  $L = 2^{6.625} = 98.7$ levels

Channel Capacity

- **Noisy Channel: Shannon Capacity**

• In reality, we cannot have a noiseless channel;
• the channel is always noisy.
• Shannon capacity is used, to determine the theoretical highest data rate for a noisy channel:

- Capacity = bandwidth * $\log_2(1 + SNR)$ bits/sec

- bandwidth is the bandwidth of the channel,
- SNR is the signal-to-noise ratio,
- capacity is the capacity of the channel in bits per second

Shannon Capacity

- Bandwidth is a fixed quantity, so it cannot be changed.
- Hence, the channel capacity is directly proportional to the power of the signal,
- as SNR = (Power of signal) / (power of noise).
- The signal-to-noise ratio (S/N) is usually expressed in decibels (dB) given by the formula:

$$10 * \log_{10}(S/N)$$

- So for example a signal-to-noise ratio of 1000 is commonly expressed as:

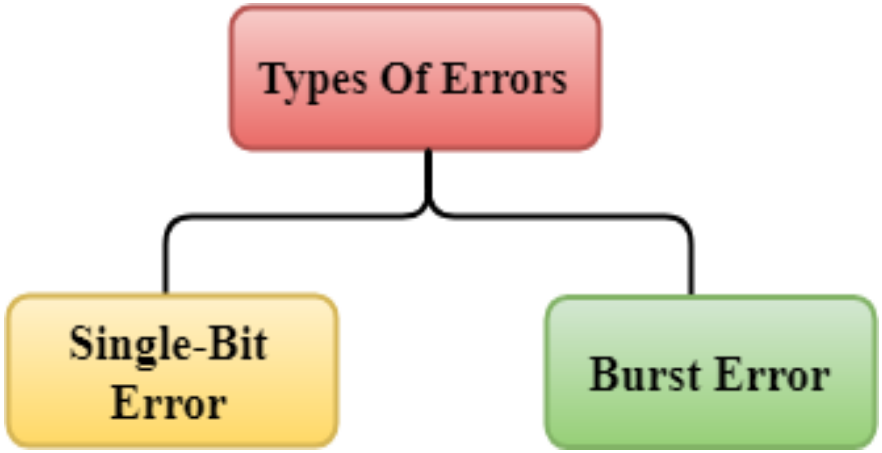
$$10 * \log_{10}(1000) = 30 \text{ dB}$$

Shannon Capacity

- **Examples:**
- **Input1:** A telephone line normally has a bandwidth of 3000 Hz (300 to 3300 Hz) assigned for data communication. The SNR is usually 3162. What will be the capacity for this channel?
  Output1 : $C = 3000 * \log_2(1 + SNR) = 3000 * 11.62 = 34860$ bps
- **Input2 :** The SNR is often given in decibels. Assume that SNR(dB) is 36 and the channel bandwidth is 2 MHz Calculate the theoretical channel capacity.
  Output2 : $SNR(dB) = 10 * \log_{10}(SNR)$
  $SNR = 10^{(SNR(dB)/10)}$
  $SNR = 10^{3.6} = 3981$
- Hence, $C = 2 * 10^6 * \log_2(3982) = 24$ MHz

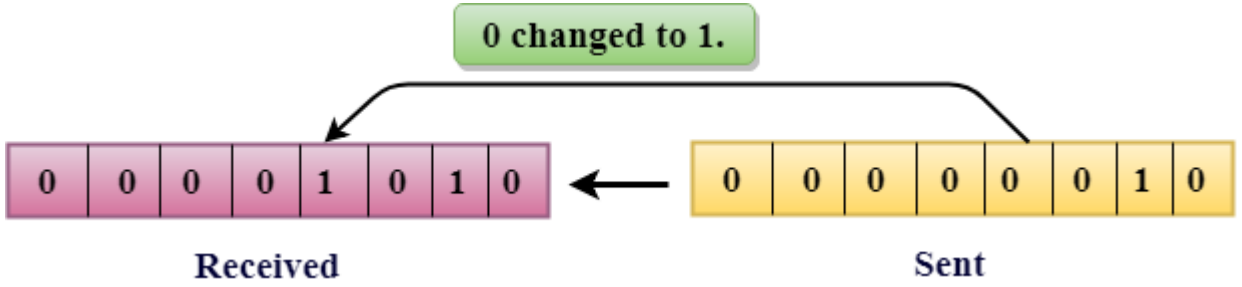
Error detection and correction

- Error is a condition when the receiver's information does not match the sender's information.
- During transmission, digital signals suffer from noise that can introduce errors in the binary bits traveling from sender to receiver.
- That means a 0 bit may change to 1 or a 1 bit may change to 0.
- **Types Of Errors**



Error detection and correction

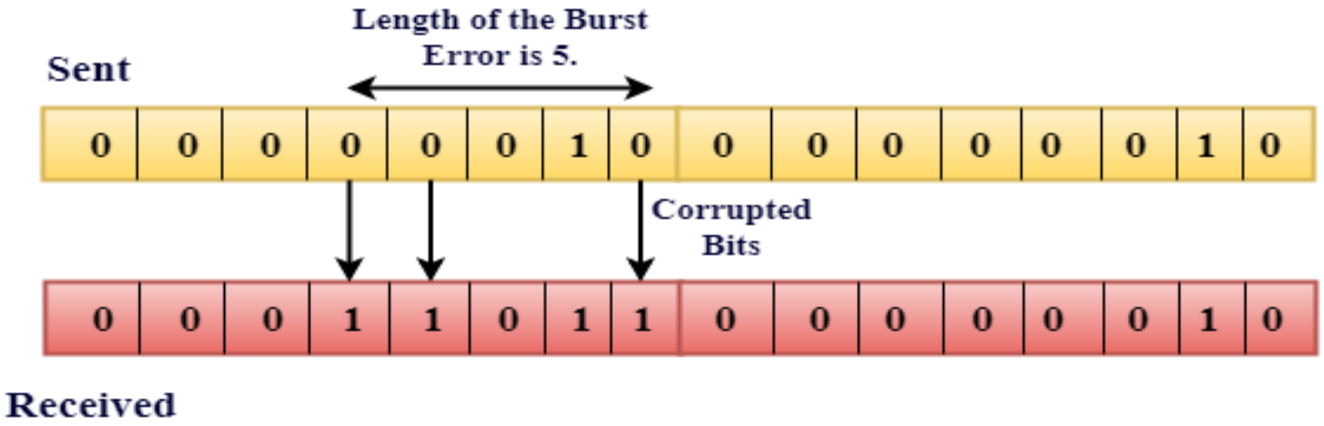
- **Single-Bit Error**
- The only one bit of a given data unit is changed from 1 to 0 or from 0 to 1.



- In the above figure, the message which is sent is corrupted as single-bit, i.e., 0 bit is changed to 1.
- **Single-Bit Error** does not appear more likely in Serial Data Transmission.
- Single-Bit Error mainly occurs in Parallel Data Transmission.
- For example, if eight wires are used to send the eight bits of a byte, if one of the wires is noisy, then single-bit is corrupted per byte.

Error detection and correction

- Burst Error
- The two or more bits are changed from 0 to 1 or from 1 to 0 is known as Burst Error.
- The Burst Error is determined from the first corrupted bit to the last corrupted bit.



- The duration of noise in Burst Error is more than the duration of noise in Single-Bit.
- Burst Errors are most likely to occur in Serial Data Transmission.
- The number of affected bits depends on the duration of the noise and data rate.

Error detection and correction

- Error control mechanism may involve two possible ways:
- Error detection
- Error correction
- **Error Detection**
- Errors in the received frames are detected by means of Parity Check and Cyclic Redundancy Check (CRC).
- In both cases, a few extra bits are sent along with actual data to confirm that bits received at the other end are the same as they were sent.
- If the counter-check at the receiver’s end fails, the bits are considered corrupted.

Error detection and correction

- The most popular Error Detecting Techniques are:
- Single parity check
- Two-dimensional parity check
- Checksum
- Cyclic redundancy check

Error detection and correction

- **Single Parity Check**
- Single Parity checking is a simple mechanism and inexpensive to detect errors.
- In this technique, a redundant bit is also known as a parity bit which is appended at the end of the data unit so that the number of 1s becomes even.
- Therefore, the total number of transmitted bits would be 9 bits.
- If the number of 1s bits is odd, then parity bit 1 is appended and if the number of 1s bits is even, then parity bit 0 is appended at the end of the data unit.
- At the receiving end, the parity bit is calculated from the received data bits and compared with the received parity bit.
- This technique generates the total number of 1s even, so it is known as even-parity checking.

Error detection and correction

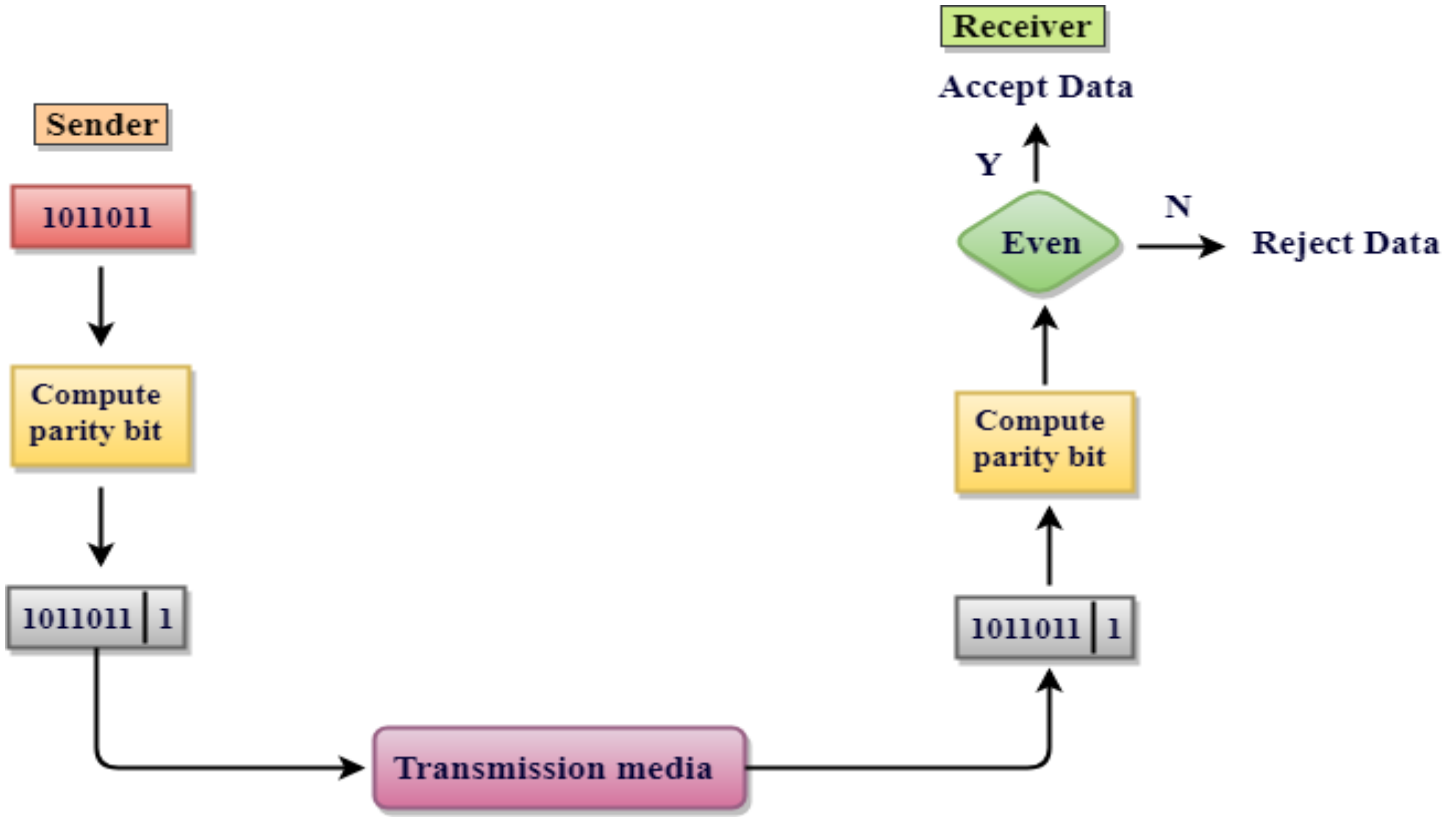
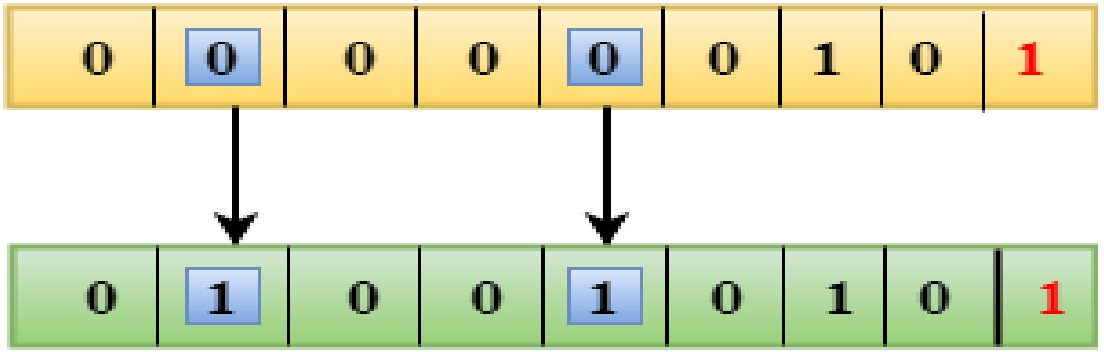


Figure : Single Parity Check

Error detection and correction

- Drawbacks Of Single Parity Checking
- It can only detect single-bit errors which are very rare.
- If two bits are interchanged, then it cannot detect the errors.



Error detection and correction

- **Two-Dimensional Parity Check**
- Performance can be improved by using the **Two-Dimensional Parity Check** which organizes the data in the form of a table.
- Parity check bits are computed for each row, which is equivalent to the single-parity check.
- In a Two-Dimensional Parity check, a block of bits is divided into rows, and the redundant row of bits is added to the whole block.
- At the receiving end, the parity bits are compared with the parity bits computed from the received data.

Error detection and correction

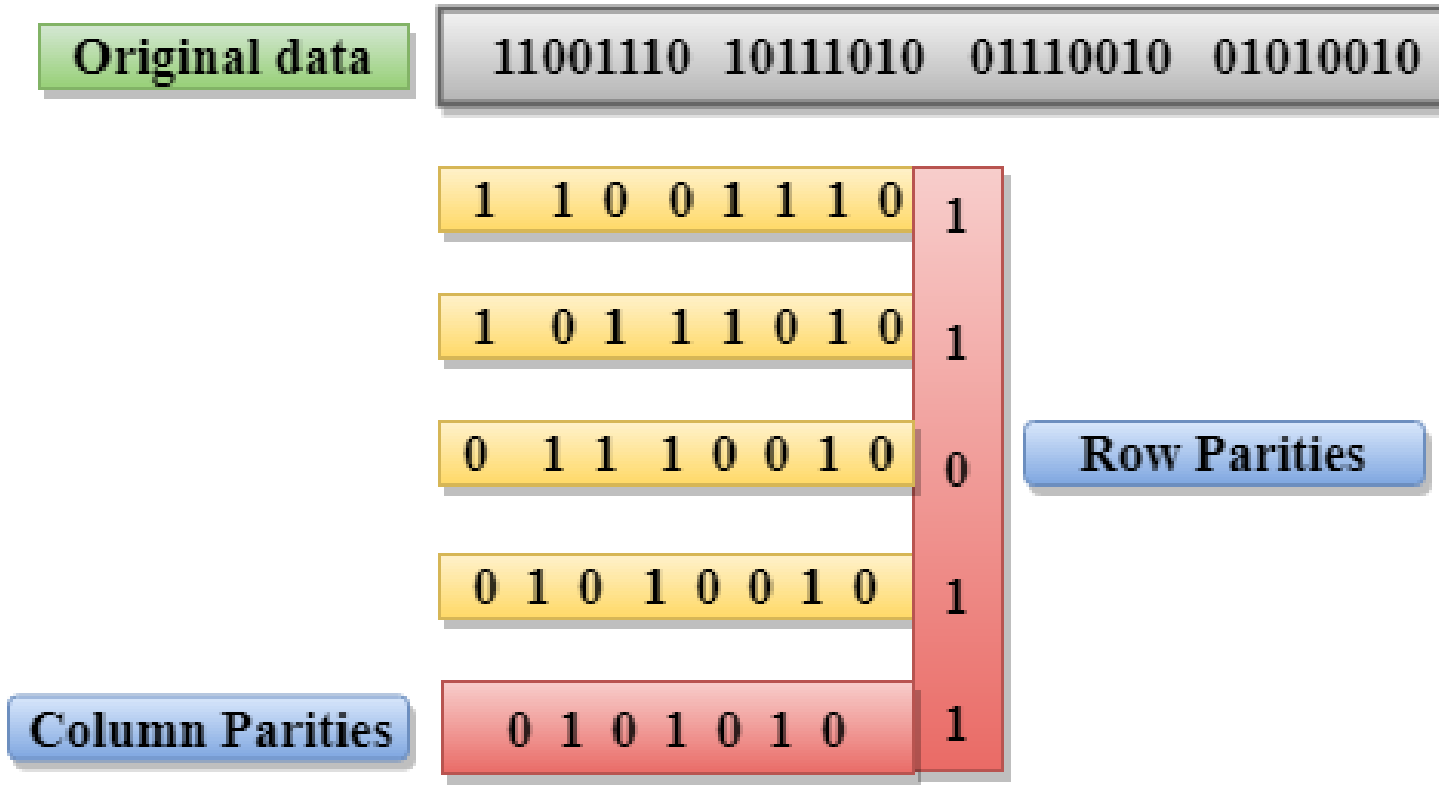


Figure : Two-dimensional Parity Check

Error detection and correction

- Drawbacks Of 2D Parity Check
- If two bits in one data unit are corrupted and two bits in exactly the same position in another data unit are also corrupted, then the 2D Parity checker will not be able to detect the error.
- This technique cannot be used to detect the 4-bit errors or more in some cases.

Error detection and correction

- **Checksum**
- A Checksum is an error detection technique based on the concept of redundancy.
- **It is divided into two parts:**
- Checksum Generator
- A Checksum is generated at the sending side.
- Checksum generator subdivides the data into equal segments of n bits each, and all these segments are added together by using one's complement arithmetic.
- The sum is complemented and appended to the original data, known as the checksum field.
- The extended data is transmitted across the network.
- Suppose L is the total sum of the data segments, then the checksum would be ?L

Error detection and correction

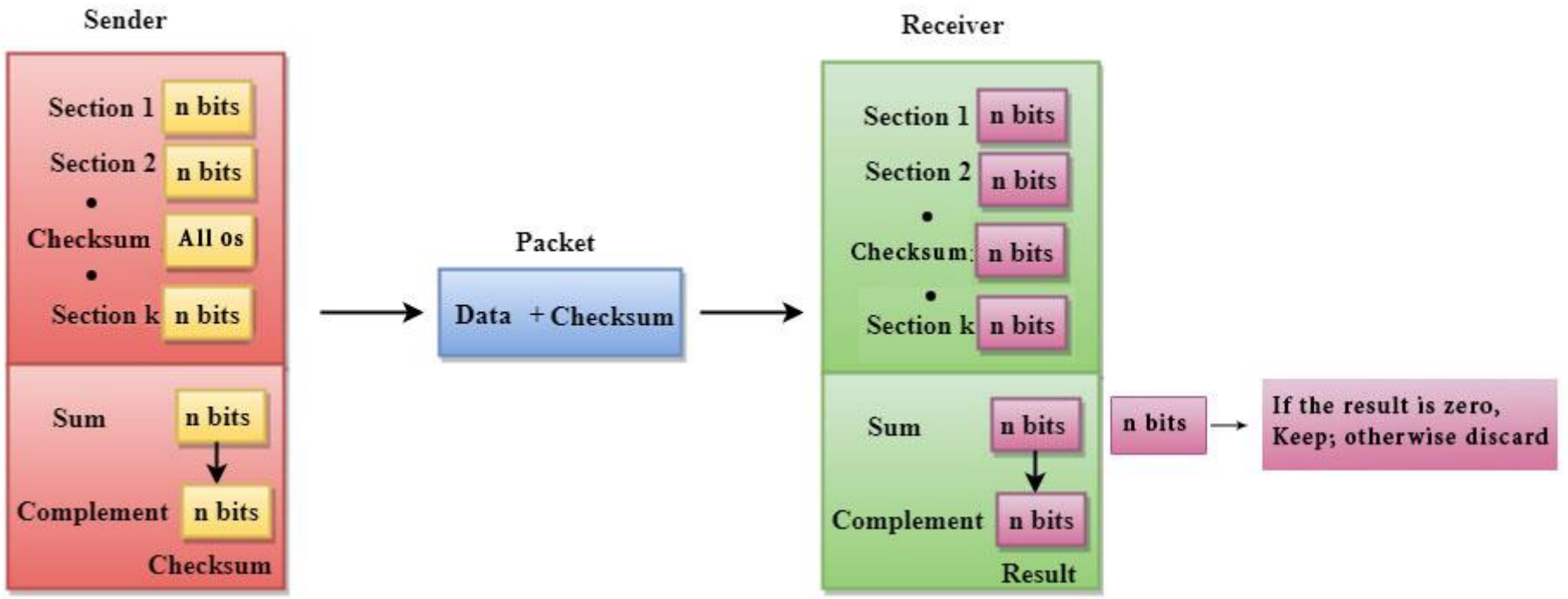


Figure: Checksum

Error detection and correction

- The Sender follows the given steps:

1. The block unit is divided into k sections, and each of n bits.
2. All the k sections are added together by using one's complement to get the sum.
3. The sum is complemented and it becomes the checksum field.
4. The original data and checksum field are sent across the network.

- **Checksum Checker**
- A Checksum is verified at the receiving side.
- The receiver subdivides the incoming data into equal segments of n bits each, and all these segments are added together, and then this sum is complemented.
- If the complement of the sum is zero, then the data is accepted otherwise data is rejected.

Error detection and correction

- The Receiver follows the given steps:

1. The block unit is divided into k sections and each of n bits.
2. All the k sections are added together by using one's complement algorithm to get the sum.
3. The sum is complemented.
4. If the result of the sum is zero, then the data is accepted otherwise the data is disc arded.

Error detection and correction

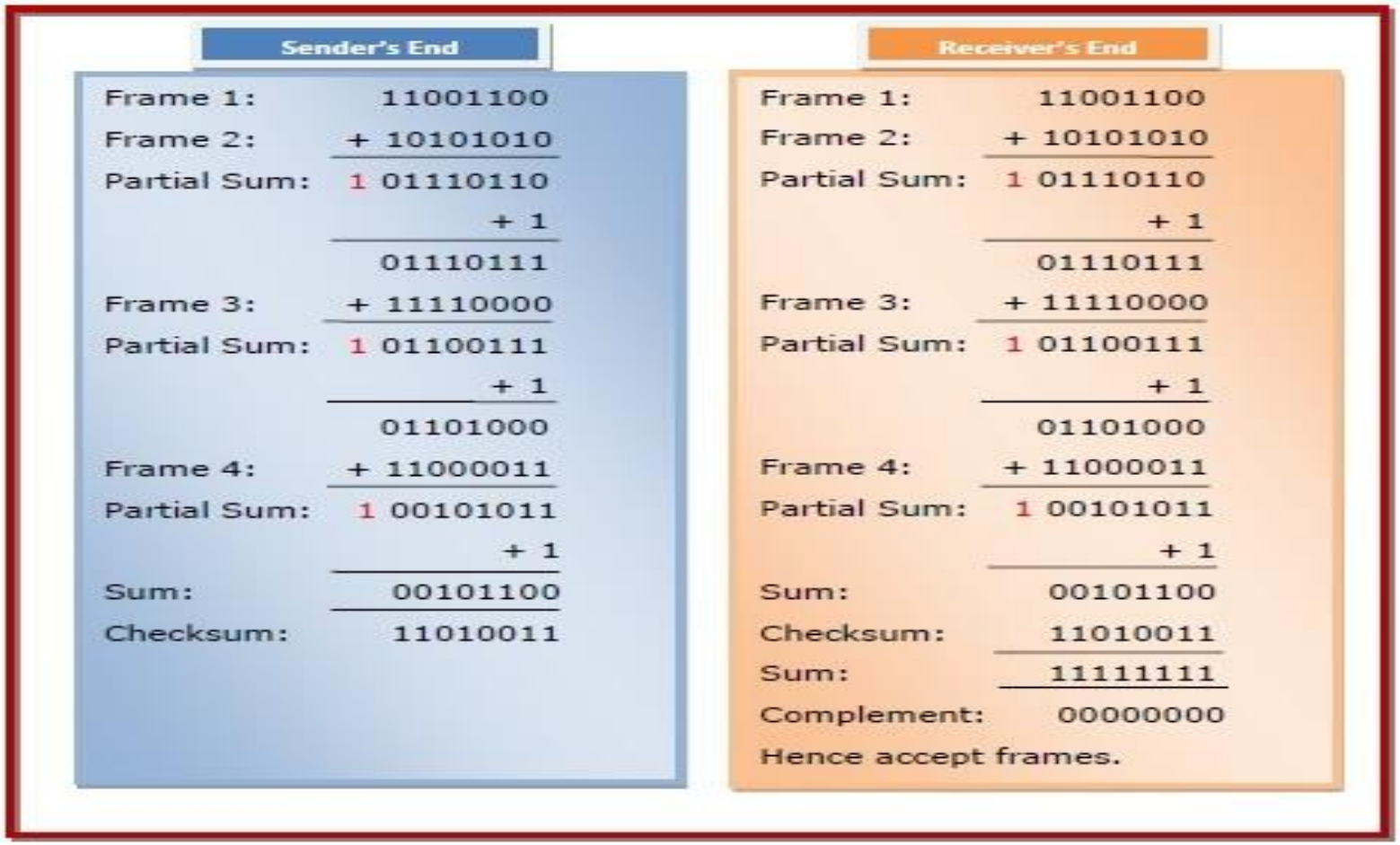


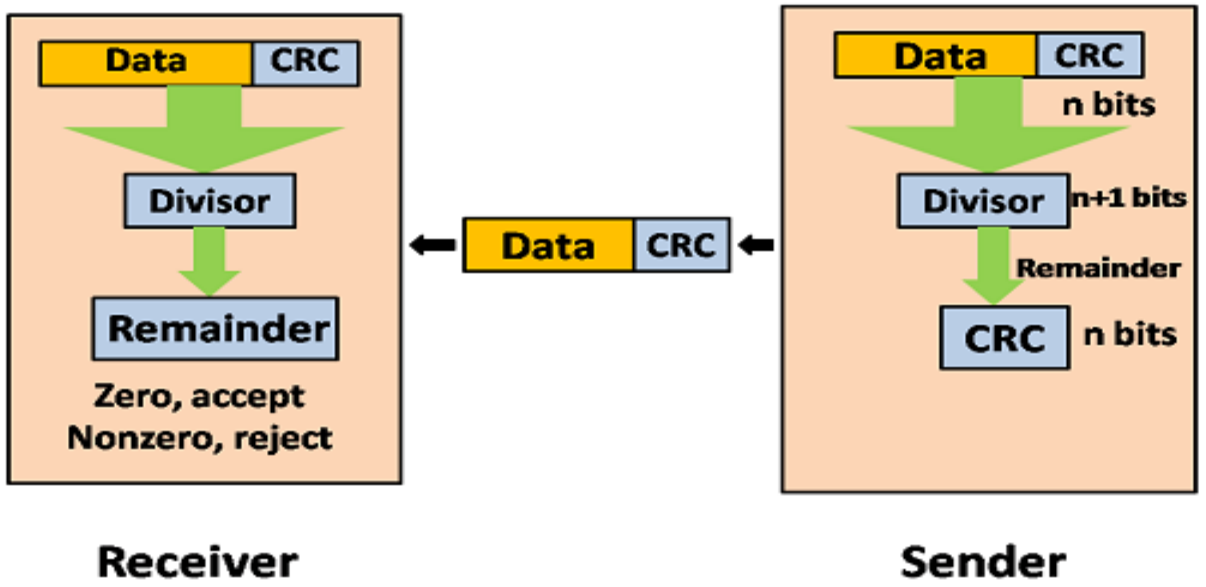
Figure: Checksum Example

Error detection and correction

- **Cyclic Redundancy Check (CRC)**
- CRC is a redundancy error technique used to determine the error.
- **Following are the steps used in CRC for error detection:**
- In the CRC technique, a string of n 0s is appended to the data unit, and this n number is less than the number of bits in a predetermined number, known as the division which is n+1 bits.
- Secondly, the newly extended data is divided by a divisor using a process known as binary division.
- The remainder generated from this division is known as CRC remainder.
- Thirdly, the CRC remainder replaces the appended 0s at the end of the original data. This newly generated unit is sent to the receiver.

Error detection and correction

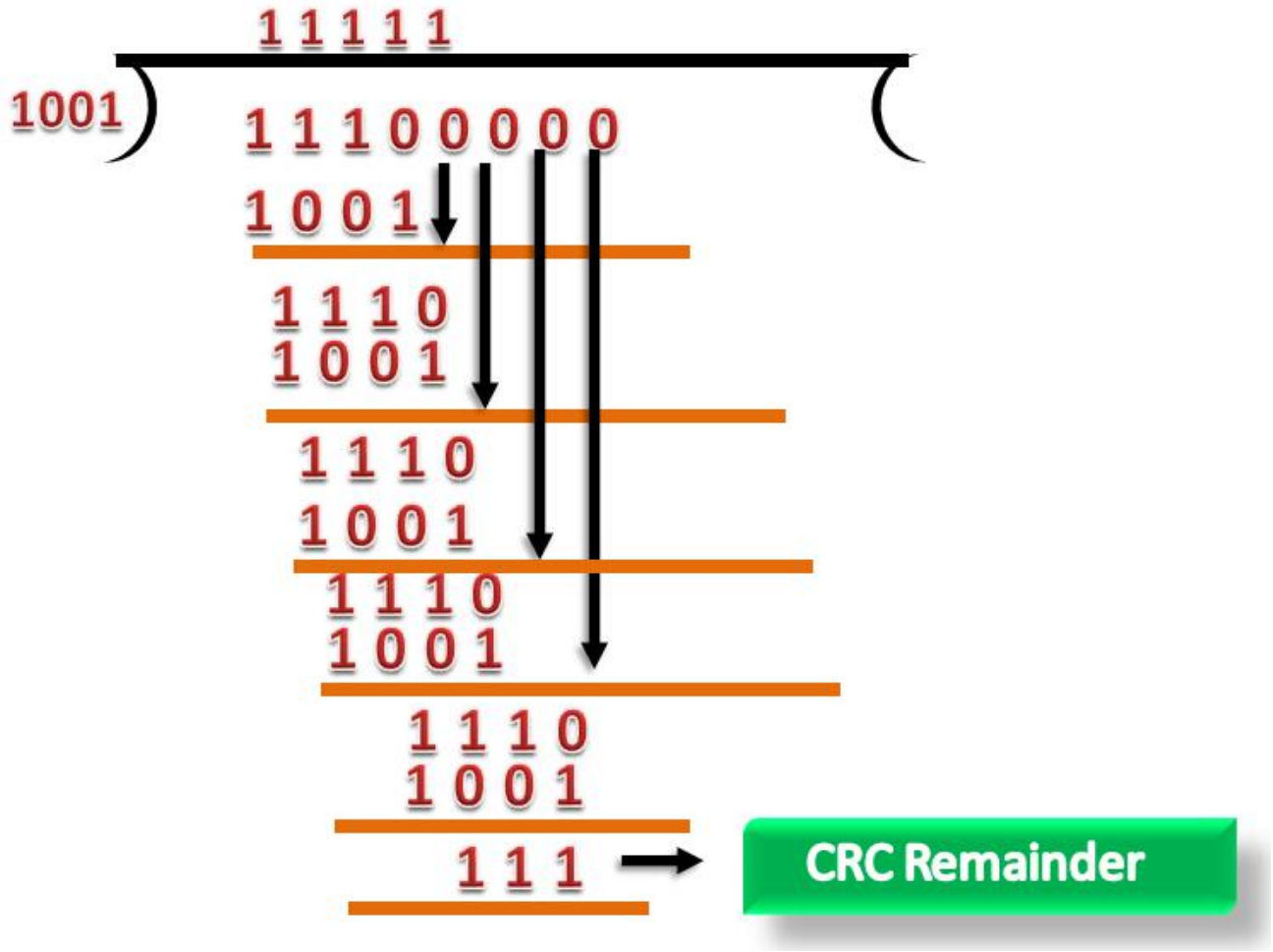
- The receiver receives the data followed by the CRC remainder.
- The receiver will treat this whole unit as a single unit, divided by the same divisor used to find the CRC remainder.
- If the resultant of this division is zero, it has no error, and the data is accepted.
- If the resultant of this division is not zero it means that the data consists of an error.
- Therefore, the data is discarded.



Error detection and correction

- **Suppose the original data is 11100 and the divisor is 1001.**
- **CRC Generator**
- A CRC generator uses a modulo-2 division.
- Firstly, three zeroes are appended at the end of the data like the length of the divisor is 4 and we know that the length of the string 0s to be appended is always one less than the length of the divisor.
- Now, the string becomes 11100000, and the resultant string is divided by the divisor 1001.
- The remainder generated from the binary division is known as CRC remainder.
- The generated value of the CRC remainder is 111.
- CRC remainder replaces the appended string of 0s at the end of the data unit, and the final string would be 11100111 which is sent across the network.

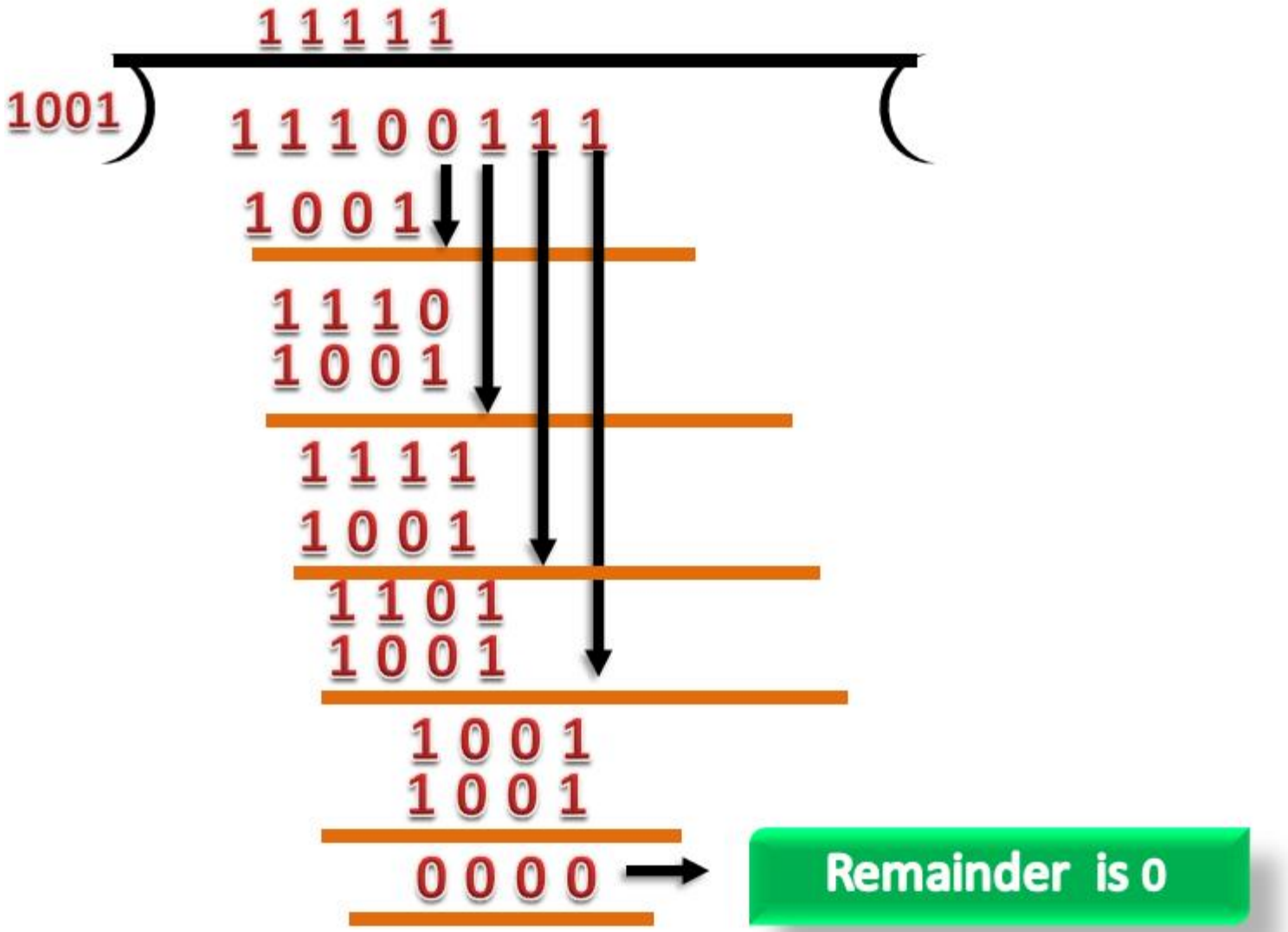
Error detection and correction



Error detection and correction

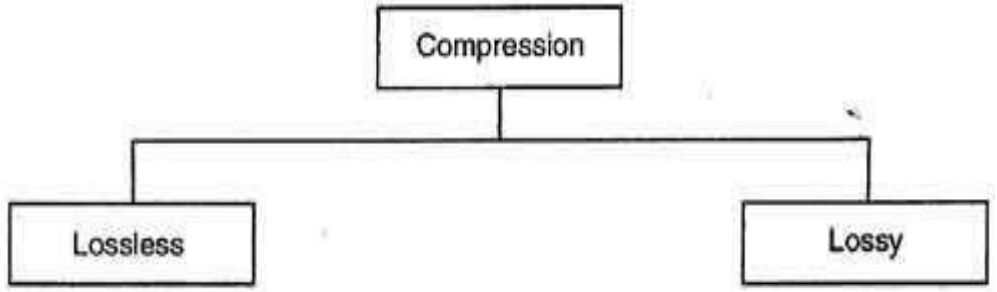
- **CRC Checker**
- The functionality of the CRC checker is similar to the CRC generator.
- When the string 11100111 is received at the receiving end, then the CRC checker performs the modulo-2 division.
- A string is divided by the same divisor, i.e., 1001.
- In this case, the CRC checker generates the remainder of zero.
- Therefore, the data is accepted.

Error detection and correction



Compression and its types

- Data compression is the function of the presentation layer in the OSI reference model.
- Compression is often used to maximize the use of bandwidth across a network or to optimize disk space when saving data.
- There are two general types of compression algorithms:



Types of Compression

Compression and its types

- **Lossless Compression**
- Lossless compression compresses the data in such a way that when data is decompressed it is exactly the same as it was before compression *i.e.* there is no loss of data.
- A lossless compression is used to compress file data such as executable code, text files, and numeric data because programs that process such file data cannot tolerate mistakes in the data.
- Lossless compression will typically not compress the file as much as lossy compression techniques and may take more processing power to accomplish the compression.
- Lossless data compression is mainly used in sensitive documents, confidential information, and PNG, RAW, GIF, and BMP file formats.

Compression and its types

- **Lossless Compression Algorithms**
- The various algorithms used to implement lossless data compression are :
- Run-length encoding
- Differential pulse code modulation
- Dictionary-based encoding
- **Run length encoding**
- This method replaces the consecutive occurrences of a given symbol with only one copy of the symbol along with a count of how many times that symbol occurs. Hence the name ‘run length’.
- For example, the string AAABBCDDDD would be encoded as 3A2BIC4D.

Compression and its types

- A real-life example where run-length encoding is quite effective is the fax machine. Most faxes are white sheets with the occasional black text.
- So, a run-length encoding scheme can take each line and transmit a code for while then the number of pixels, then the code for black and the number of pixels, and so on.
- This method of compression must be used carefully.
- If there is not a lot of repetition in the data then it is possible the run-length encoding scheme would actually increase the size of a file.

Compression and its types

- **Differential pulse code modulation**
- In this method first a reference symbol is placed.
- Then for each symbol in the data, we place the difference between that symbol and the reference symbol used.
- For example, using symbol A as the reference symbol, the string AAABBC DDDD would be encoded as AOOOl123333, since A is the same as the reference symbol, B has a difference of 1 from the reference symbol, and so on.
- **Dictionary based encoding**
- One of the best-known dictionary-based encoding algorithms is Lempel-Ziv (LZ) compression algorithm.
- This method is also known as a substitution coder.

Compression and its types

- In this method, a dictionary (table) of variable length strings (common phrases) is built.
- This dictionary contains almost every string that is expected to occur in data.
- When any of these strings occur in the data, then they are replaced with the corresponding index to the dictionary.
- In this method, instead of working with individual characters in text data, we treat each word as a string and output the index in the dictionary for that word.
- For example, let us say that the word “compression” has the index of 4978 in one particular dictionary; it is the 4978^{th} word is usr/share/dict/words.
- To compress a body of text, each time the string “compression” appears, it would be replaced by 4978.

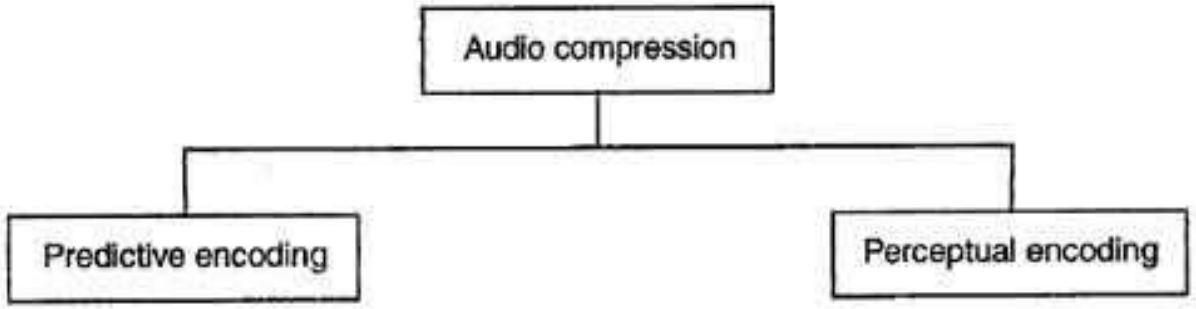
Compression and its types

- **Lossy Compression**
- Lossy compression is the one that does not promise that the data received is exactly the same as the data send *i.e.* the data may be lost.
- This is because a lossy algorithm removes information that it cannot later restore.
- Lossy algorithms are used to compress still images, video, and audio.
- Lossy algorithms typically achieve much better compression ratios than lossless algorithms.
- Lossy data compression is most widely used in JPEG images, MPEG video, and MP3 audio formats.

- **Audio Compression**
- Audio compression is used for speech or music.
- For speech, we need to compress a 64-kHz digitized signal; For music, we need to compress a 1.411.MHz signal

Compression and its types

- Two types of techniques are used for audio compression:



Techniques of audio compression

Compression and its types

- **Predictive encoding**
- In predictive encoding, the differences between the samples are encoded instead of encoding all the sampled values.
- This type of compression is normally used for speech.
- Several standards have been defined such as GSM (13 kbps), G. 729 (8 kbps), and G.723.3 (6.4 or 5.3 kbps).
- **Perceptual encoding**
- Perceptual encoding scheme is used to create CD-quality audio that requires a transmission bandwidth of 1.411 Mbps.
- MP3 (MPEG audio layer 3), a part of MPEG standard uses this perceptual encoding.
- Perceptual encoding is based on the science of psychoacoustics, a study of how people perceive sound.

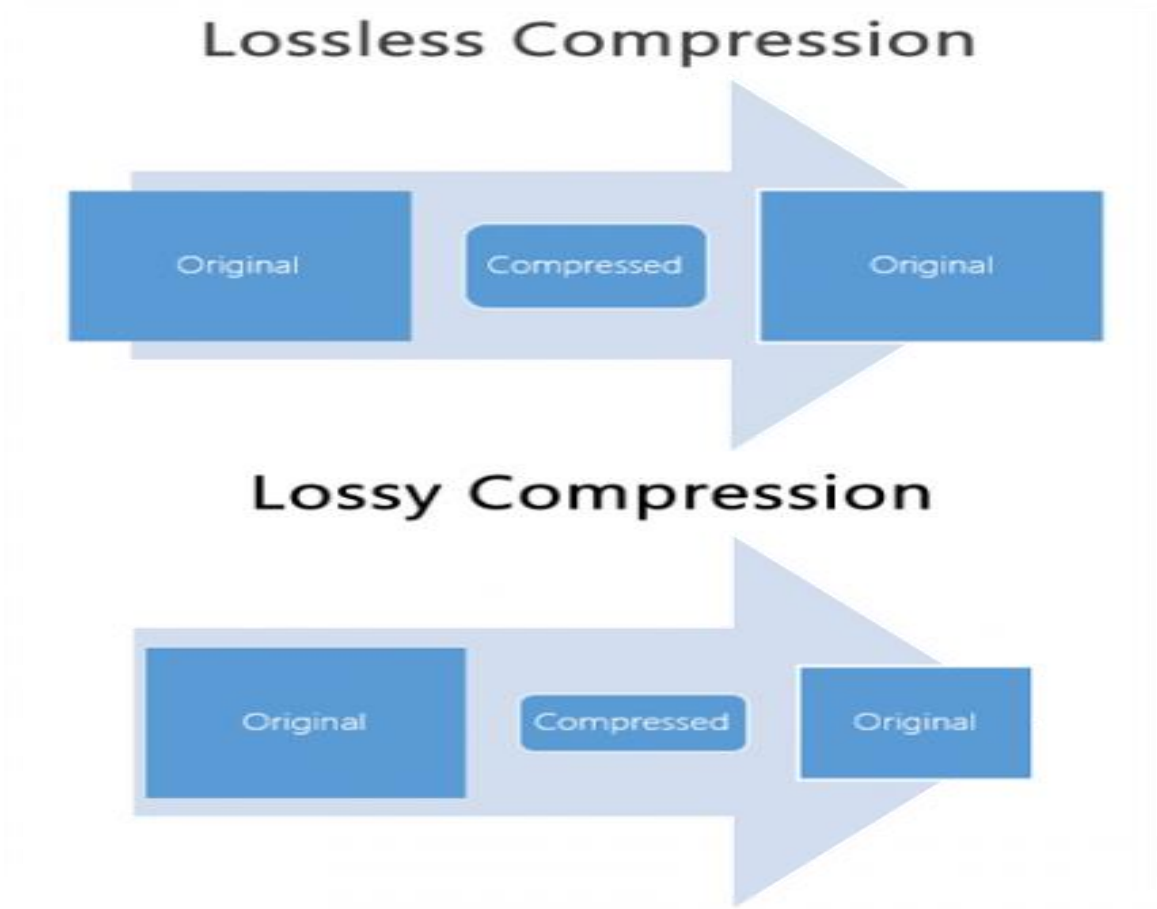
Compression and its types

- The perceptual encoding exploits certain flaws in the human auditory system to encode a signal in such a way that it sounds the same to a human listener, even if it looks quite different on an oscilloscope.
- The key property of perceptual coding is that some sounds can mask another sound.
- For example, imagine that you are broadcasting a live flute concert and all of a sudden someone starts striking a hammer on a metal sheet.
- You will not be able to hear the flute anymore. Its sound has been masked by the hammer.

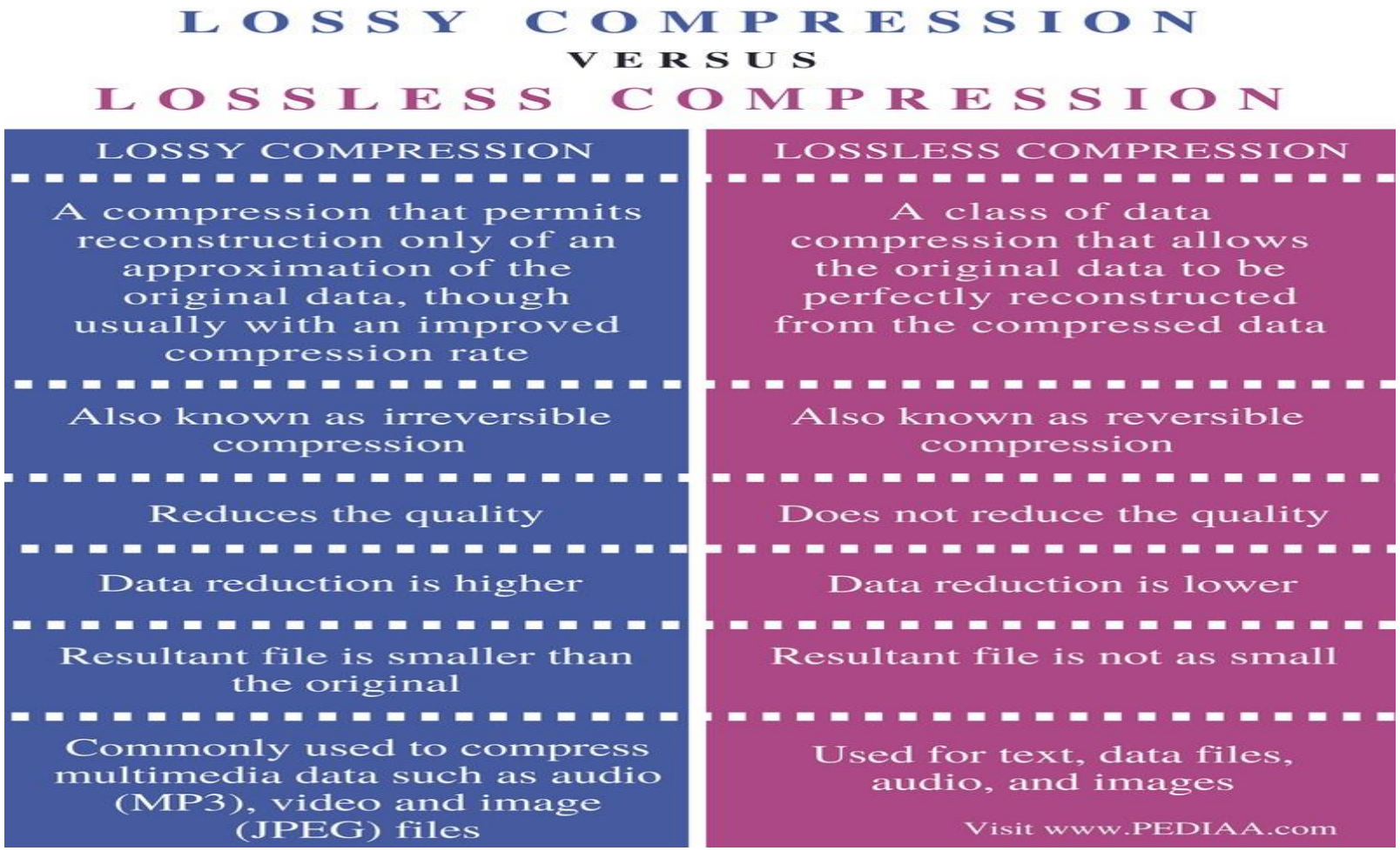
Compression and its types

- **MP3**
- MP3 uses these two phenomena, *i.e.* frequency masking and temporal masking to compress audio signals.
- In such a system, the technique analyzes and divides the spectrum into several groups.
- Zero bits are allocated to the frequency ranges that are totally masked.
- A small number of bits are allocated to the frequency ranges that are partially masked.
- A larger number. of bits are allocated to the frequency ranges that are not masked.
- Based on the range of frequencies in the original analog audio, MP3 produces three data rates: 96kbps, 128 kbps, and 160 kbps.

Compression and its types



Compression and its types



THANK YOU