**Data Compression**

Kompresimi i të dhënave ndodh për tri arsye kryesore : 1.Për të kursyer hapsirën kur të dhënat ruhet, 2.Për të kursyer kohën kur të dhënat transmetohen, 3.Sepse shumica e fajllave(të dhënave) kanë shumë përsëritje.

Kur të dhënat kompresohen ato në të njëjtën kohë edhe enkriptohen gjë që siguron të dhënat.

Pyetjes se për çfarë nevojitet kompresimi i të dhënave i përgjigjen dy ligje aj i Moor-it dhe aj i Parkisonit. Ligji i Moorit thotë se numri tranzistorëve nëpër çipa dyfishohet gjdo 18-24 muaj, kjo na tregon se hapsira memoruese rritet shumë. Ligji i Parkisonit thotë se të dhënat shtohen aq sa ka hapsirë për tu mbuluar. Nga kjo konkludojmë se gjdo dit numri të dhënave rritet masivishtë , për ta bër menagjimin më të let të tyre kto të dhëna duhet të kompresohen.

Kompresimi i të dhënave ka filluar që nga vitet e 50-ta , ku për qëllim ka qenë kursimi i memories, por në ditët e sotme qëllim kryesorë është kursimi i kohës transmetuese.

Aplikimet e Kompersimit : **në fajlla gjenerik** : GZIP, 7Z, NTFS, ZFS, **në multimedia** : GIF, JPEG, MP3 **, në komimunikim** : skype , **në databaza** etj.

Modeli bazik i kompresimit :

Teorikisht komopresimi i të dhënave ndodhë pa humbjen e të dhënave, mirpo në parim nuk është e mundur një gjë e tillë sepse pa mar parasysh sa i fuqishëm është algoritmi së paku një bit do humbet gjatë zgjerimit. Kjo përshkak të materialeve përquese.

Përveq në kompjuterik , kompresimi ndodhë po thuajse gjithkund duke filluar nga truri i njeriut , gjuhët , sistemet telefonike , sistemet numerike , moorse code etj.

Kompresimi i të dhënave ndikon në shpejtimin e algoritmeve të kërkimit(search).

Fajllat të cilët nuk kanë përsëritje nuk preferohet dhe nuk duhet të kompresohen sepse vetëm sa e rrisin madhësinë e fajllit.

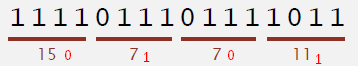
Algoritmet e kompresimit ndahen në **tri modele** bazë : **1.Statike , 2.Dinamike , 3.Adaptive**.

Algoritmet të bazuara në modelin **statik** : Përdoret modeli i njëjtë për të gjitha tekstet, është i shpejtë , por nuk është optimal sepse tekstet e ndryshme kanë veçi të ndryshme statistikore ,shembuj : Run-Length code, ASCII, Morse Code.

Algoritmet të bazuara në modelin **dinamik** : Modeli gjenerohet duke u bazuar në tekst, ky model duhet të transmetohet së bashku me informacionin e kompresuar, shumebull : Huffman code.

Algoritmet të bazuara në modelin **adaptiv** : Modeli mësohet dhe përditsohet np mënyrë progresive për gjatë leximit të tekstit. Modelimi më i saktë jepë kompresim më të mirë, dekodimi starton nga fillimi, shembull : LZW.

**Run-Length Codding** :Përdoret kur kemi radhë të gjatë të bitave të cilët përsëriten, kjo teknikë e kompresimit nuk punon kur kemi fajlla pa përsëritje vetëm se e rrit madhësinë e fajllit. Kjo teknikë përdoret nga JPEG, Fax, etj.



**Haffman Commpresion :** Bazohet në modelin dinamik , idea e përdorimit të variable-length code vjen sepse nganjëher është problem të dekodohet kodi me gjatësi fikse. Ky algoritëm është dinamik sepse , së pari e lexon inputin pas leximit i mer frekuencat e të gjitha karaktereve dhe në baz të tyre gjeneronë kodin. Kjo gjë edhe na qon drejt të metës kryesore të kti algoritmi pra i nevojitet fillimisht leximi i inputit kjo në transmetim shkakton vones shumë të madhe. Po përsëri për aplikime lokale është një algoritmë shumë i fuqishëm. Ky algoritëm implemetohet me anë të tires.

**LWZ :** Ky algoritëm bazohet në modelin adaptiv dhe gjen zbatim të madhë në Real-Time aplikacione si ato për komunikim , gaming etj e met e kti algoritmi është se kur ka përsëritje shumë procesisi i dekodimit mund të bphet pak tricky.Por kjo gjë ndodhë rradhë.