

LA COMPRESSION DES DONNEES

I- Généralité :

1- Définition :

Réduire les données des fichiers informatiques pour en stocker davantage ou les transférer plus rapidement sans les altérer (ou du moins pas trop) est devenue une opération de routine indispensable. Car sans cette opération de compression de données, il n'y aurait pas d'Internet à haut débit, ni de lecteurs MP3, ni de télévision satellite, ni de téléphones portables. La liste des services que rend tous les jours la compression des données est interminable.

2- Origine :

Comme très souvent dans les domaines des sciences et des techniques, de nombreuses personnes ont participé et participent aux divers développements de la compression des données. Mais notre société informatisée, très friande de compression, doit beaucoup à la **théorie de l'information de Claude Shannon**. Cet ingénieur américain était employé par le département de cryptographie des services secrets de son pays durant la Seconde Guerre mondiale. Son rôle y était de décortiquer le code secret très brouillé de l'armée allemande et d'y trouver de l'information exploitable. Une fois la guerre terminée, cette expérience lui a permis d'écrire un livre dans lequel il exposait sa théorie de l'information.

3- Méthode et Risques :

On parle d'elle généralement au singulier, mais il existe plusieurs types de compression des données qui reposent sur différents **algorithmes** (procédés de calcul informatique). Le principe de base est le même : trouver dans l'élément à compresser des zones identiques ou inutiles pour les simplifier ou les supprimer. C'est l'élément à compresser qui va déterminer le type de compression à appliquer.

Quand un fichier est compressé avec pertes, cela signifie qu'on le dégrade pour l'alléger. Le programme de compression recherche ce qu'il est possible de supprimer sans rendre le fichier illisible.

Par exemple, si l'on compresse correctement un **fichier audio**, les hautes et les basses fréquences considérées comme inaudibles pour l'oreille humaine sont supprimées. Dans ce cas, la musique reste agréable à écouter et seule une oreille très avertie pourra parfois remarquer la perte de fréquences.

Si l'on souhaite alléger encore plus le fichier audio, on peut choisir une compression plus importante qui supprimera alors des fréquences audibles. Ce type de fichiers audio reste écoutable, mais perd en qualité.

On peut également compresser les **fichiers vidéo**. Un fichier vidéo est composé d'une suite d'images souvent partiellement identiques. Dans une scène où deux personnages discutent assis à une table, plusieurs éléments ne bougent pas d'une image à l'autre : la table, les chaises, les verres posés sur la table, etc. Il est donc inutile de les reproduire à chaque image et seul ce qui est en mouvement est pris en compte et recalculé.

Pour les **fichiers images**, l'algorithme appliqué est différent. La compression d'image fixe consiste à trouver les nuances d'une même couleur qui se trouvent à proximité dans l'image pour n'en faire plus qu'une et réduire ainsi la taille du fichier. Il en résulte une perte de détails mais lorsque la compression n'est pas trop importante, l'œil humain s'en accommode parfaitement.

La compression sans pertes est utilisée pour les données qui doivent rester identiques à l'original une fois décompressée. **Elle est donc réversible.**

Elle a deux principales applications qui sont la **transmission** et le **stockage**. Les réseaux de communication (Internet, téléphones mobiles, transmissions satellite, etc.) seraient vite saturés si on ne compressait pas les données avant de les transmettre. Quand on parle dans un téléphone mobile, avant que la voix ne parvienne à un interlocuteur, elle subit divers traitements : le téléphone la numérise (elle est transformée en une suite de 0 et de 1), la compresse, la découpe en petits paquets qu'il envoie sur le réseau téléphonique ; lorsqu'elle arrive sur l'autre téléphone, celui-ci regroupe les petits paquets, les décompresse et transforme la suite de 0 et de 1 en sons audibles par l'oreille humaine. Il en va de même pour la majorité des fichiers qui circulent sur Internet.

Pour effectuer une compression sans pertes, le programme recherche les éléments qui se répètent pour pouvoir les regrouper et les simplifier.

Qu'est-ce que le taux de compression ?

Il s'agit du **pourcentage de compression appliqué à un fichier**. Pour la compression avec pertes, le taux de compression doit être choisi avec attention pour ne pas rendre le fichier inutilisable. En revanche, la compression sans pertes peut supporter des taux de compression très élevés.

Il faut également noter que plus le taux est élevé, plus il faut de temps pour les opérations de compression et de décompression.