

**VISUALISATION D'INFORMATION**  
**Durée 2h – 2 pages**

Tous documents manuscrits autorisés.

Notes dactylographiées distribuées en Cours autorisées.

**Ce sujet est un sujet de réflexion : il n'existe pas de solution unique. Argumentez brièvement mais soigneusement vos choix.**

Dans tout ce sujet, on s'intéresse à la visualisation d'un système de fichiers partagés que l'on nommera SGF.

1 – Donnez les raisons pour lesquelles un SGF peut être représenté comme un graphe. Quelles sont les propriétés de ce graphe ?

2 – Donnez des exemples de sous arbres recouvrant pouvant représenter le point de vue de l'utilisateur dans chacun des cas suivants :

- Un utilisateur de base,
- Un administrateur du SGF,
- Un ingénieur chargé d'installer des binaires.

3 – On suppose que l'on possède la liste des nombres de visites effectuées par jour sur chaque fichier durant une période de  $k$  journées.

- Quels types de représentations graphiques peut-on proposer pour mettre en évidence les sommets les plus visités pour une journée fixée? On précisera les attributs de la représentation et éventuellement les modifications à apporter aux représentations usuelles.
- Que faire si on change « plus visité » en « moins visité »
- On désire mettre en évidence les fichiers les plus visités durant les  $k$  journées. Quelle métrique peut-on construire? Quel algorithme peut-on envisager d'utiliser pour partitionner le graphe en utilisant la métrique « nombre de visites »? Ce mécanisme peut-il s'appliquer aux sous arbres décrits au point 2.

4 – La personne chargée d'administrer le SGF souhaite mettre en relief des consommations de ressources disques abusives ayant les critères suivants : taille, MPEG3, JPEG. Quelles lentilles magiques pouvez-vous proposer, dans les cas suivant :

- Pour les développeurs d'une société fabriquant des jeux vidéos,
- Pour des statisticiens,
- Pour le personnel non développeur d'une société n'ayant aucune vocation audio ou image.

5 – On suppose que le système est capable de tracer les copies de documents entre des répertoires utilisateurs différents.

- Quel nouveau graphe peut-on considérer ?
- Ce graphe a-t-il des propriétés particulières ?
- Quel algorithme de dessin de ce graphe pouvait vous proposer ?
- En fonction de l'algorithme choisi, donnez une interprétation de la vue.

6 – On souhaite modéliser les SGF afin de les générer aléatoirement.

- Quel est l'intérêt de disposer d'un générateur aléatoire ?
- Supposons que les utilisateurs n'utilisent jamais la commande unix « ln », quel modèle peut-on proposer pour les SGF des utilisateurs ? Quel type de génération aléatoire pouvez-vous envisager ?
- L'analyse statistique a montré que dans la partie système les liens symboliques représentent 5% du nombre de fichier. Comment modifier l'algorithme de génération aléatoire de DAG pour obtenir un graphe aléatoire pouvant représenter la partie système du SGF ? On ne se posera pas le problème de convergence.
- Quelles données supplémentaires doit on connaître pour générer aléatoirement un SGF complet ? On s'inspirera de l'algorithme de génération aléatoire de Web graphes.