

**VII. Conclusion :**

Bien que le modèle de sécurité utilisé est simple (semi - honnête), les protocoles proposés ne fournissent pas une préservation complète de privacy dans l'algorithme k-means, Le problème majeur est dans les itérations de l'algorithme où les informations à protéger sont nombreuses : les distances, les affectations aux clusters, le nombre de items dans chaque cluster, les centroïdes intermédiaires et le nombre des itérations et la protection des items elles même. Le modèle de distribution de données influe aussi sur la façon dont la privacy est préservée. Nous pensons que le modèle de partitionnement arbitraire d'un ensemble de données est le mieux adapté, pour chercher une solution plus générale. Cependant la plupart des travaux dans ce modèle sont appliqués sur deux parties seulement sauf que dans [104]. Tous les travaux sur le modèle vertical sont donnés sur des parties multiples mais en considérant des parties de confiance non concertées [95][94], où aucune garantie de sécurité n'est donnée si ces parties deviennent concertées.

Le coût de privacy est mesuré par rapport aux coûts de de communication du protocole. L'idéal est d'avoir un protocole optimal, permettant de s'appliquer sur un large ensemble de données avec un temps de calcul et de communication restreint, ceci revient principalement aux primitives de privacy utilisées. Les protocoles basés sur le circuit d'évaluation de Yao [6] sont très coûteux par apport à ceux basés sur les crypto systèmes homomorphes [103], mais même le coût de chiffrement dans ces derniers n'est pas négligeable. Les schémas de secret partagé additif sont prometteurs car ils présentent un coût minimal de calcul, même si leur utilisation prévoit des parties de confiance non concertées [95]. Cependant, la majorité de ces travaux appliquent le protocole de Yao avec les cryptos systèmes homomorphes.

Cette étude nous a permis de ressortir les besoins de préservation de privacy dans chaque modèle de distribution de données pour l'algorithme de clustering k-means, en mesurant la protection des items dans les étapes de l'algorithme suivant le modèle de sécurité semi - honnête. L'intérêt est aussi de ressortir les meilleurs travaux de préservation de privacy dans l'algorithme k-means en termes de préservation de privacy.