



HAL
open science

Allocation Equitable des Tâches de Collecte Participative de Données

Rim Ben Messaoud, Yacine Ghamri-Doudane

► **To cite this version:**

Rim Ben Messaoud, Yacine Ghamri-Doudane. Allocation Equitable des Tâches de Collecte Participative de Données. CoRes 2016, May 2016, Bayonne, France. hal-01306465

HAL Id: hal-01306465

<https://hal.science/hal-01306465>

Submitted on 25 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Allocation Équitable des Tâches de Collecte Participative de Données

Rim Ben Messaoud^{*†} and Yacine Ghamri-Doudane[†]

[†] *Université de la Rochelle, La Rochelle, France*

^{*} *Université Paris-Est Marne La Vallée, Champs-sur-Marne, France*

La collecte participative de données est un paradigme émergent qui permet aux utilisateurs munis de dispositifs intelligents de collecter et de partager des données sur un phénomène particulier. Pour assurer une bonne qualité d'information (QoI), il faut garantir l'engagement des participants tout en respectant la contrainte énergétique de leurs mobiles. Ce trade-off est difficile à réaliser vu la nature contradictoire des deux objectifs de la maximisation de la QoI et la minimisation de la fréquence de capture de données des terminaux des utilisateurs. Dans cet article, nous élaborons un modèle d'allocation équitable des tâches de collecte participative afin d'optimiser l'autonomie des terminaux et atteindre le niveau de QoI requis par les services/usagers demandeurs. Nous présentons d'abord le problème d'optimisation multi-objectif correspondant. Ensuite, nous proposons l'algorithme d'allocation équitable des tâches de sensing, (F-QEMSS), où nous avons eu recours à l'heuristique de la recherche taboue (TS) pour résoudre ce problème. Les résultats de la simulation montrent l'efficacité de notre solution. Particulièrement, F-QEMSS réalise un taux d'équité de 96% mesuré par l'indice de Jain tout en préservant le même niveau de QoI atteint par des algorithmes d'allocation non équitables.

Keywords: Participatory-Sensing, Allocation Équitable, QoI, Énergie

1 Introduction

Les nouveaux terminaux mobiles sont dotés d'une variété de capteurs embarqués tels que l'accéléromètre, le GPS, la caméra et le microphone en plus des capacités de calcul et de communication. Ces dernières appuient leur rôle dans la collecte participative et le partage des données, définis comme Participatory-Sensing. Plusieurs applications reposant sur ce modèle de collecte ont été proposées dans la littérature comme la gestion du trafic routier et la protection de l'environnement [1].

Toutefois, un tel paradigme prometteur dépend essentiellement de l'engagement des utilisateurs. En fait, les participants exploitent leurs ressources énergétiques afin d'exécuter des campagnes de collecte et peuvent être activement impliqués dans certaines tâches comme la prise de photos. Ainsi, ils devraient être recrutés sur la base des modèles de récompense, dits "incentives" ou participer volontairement. Dans les deux cas, une attribution adéquate des tâches s'avère nécessaire pour répondre aux exigences de qualité de données des demandeurs et garantir aux participants un processus de collecte équitable et efficace.

Dans ce contexte, des travaux de recherche ont étudié l'optimisation de la collecte participative en termes de minimisation du coût énergétique [2] ou des incitations. En outre, nous avons proposé, dans un travail précédent [3], un modèle d'allocation des tâches de sensing qui vise à maximiser la qualité des données acquises tout en minimisant la consommation d'énergie globale. Tous ces modèles se basent sur la collaboration entre les participants mais aucun n'a examiné l'équité de l'allocation finale.

D'autre part, les auteurs de [4] et [2] ont étudié respectivement l'équité lors du partage des récompenses ou du nombre de tâches affectées. Cela a été élaboré d'une façon qualitative et sans considérer la qualité des données atteinte, une exigence majeure du *Participatory Sensing*. Ainsi, nous avons pour but dans ce travail de maximiser la QoI dans une attribution de tâche équitable et respectueuse de la contrainte d'énergie. Nous avons également évalué les performances de notre système par différentes métriques d'équité.

Nous présentons dans un premier temps notre scénario de collecte participative puis sa modélisation en un problème d'optimisation multi-objectif. Nous détaillons ensuite l'exploitation de l'algorithme de la recherche taboue pour l'implémentation de notre méthode, F-QEMSS. Les résultats de simulation sont illustrés aussi tout en comparant notre modèle de déploiement par rapport aux méthodes de l'état de l'art.

2 Modélisation du système

2.1 Système proposé

Nous considérons un ensemble de N participants et un serveur centralisé dans le cloud comme l'indique la Figure 1. Chaque utilisateur, muni de son dispositif mobile équipé de capteurs embarqués, est considéré comme un candidat au Crowdsensing. La mobilité des participants est supposée connue et peut être déduite de l'historique de leurs traces. Au début de chaque période, le serveur notifie les participants des futures campagnes de collecte. Chaque participant met à jour sa position et le niveau de batterie de son mobile. En se basant sur ces données, le serveur détermine le sous-ensemble de participants qui assure un maximum de QoI et d'équité en termes de ressources dédiées.

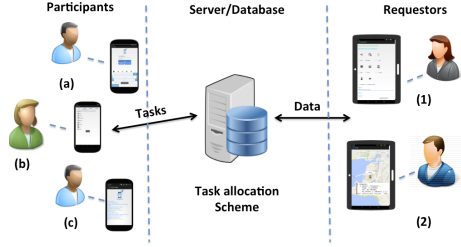


FIGURE 1: Les participants collectent et transmettent les données requises par des tiers.

2.2 Formulation du problème d'optimisation

Le but de ce travail est de trouver le sous-ensemble d'utilisateurs qui réalise un maximum de QoI des données collectées et d'équité tout en minimisant la consommation des ressources dédiées. Pour ce faire, nous précisons d'abord les attributs de qualité de données à considérer dans notre scénario ; la précision temporelle, la complétude et l'accessibilité des données. Le premier terme désigne le fait qu'une mesure doit être collectée dans un intervalle de temps précis. Le deuxième incite à couvrir le maximum d'espace. Enfin, nous considérons l'accessibilité d'une donnée en termes d'énergie nécessaire pour la collecter. Ces attributs sont quantifiés et normalisés, comme dans [3], par une fonction d'utilité proposée dans [5]. Pour chaque attribut nous associons une forme de fonction d'utilité qui lui correspond et que nous détaillons dans notre travail [3]. De plus, nous proposons de mesurer l'équité de notre solution par l'indice de Jain défini par : $J = (\sum_{i=1}^N \delta_i)^2 / p \sum_{i=1}^N \delta_i^2$. Où p est le nombre des participants sélectionnés pour la collecte et $\delta_i = U(n_i)$ est le nombre des tâches de sensing normalisé par la même fonction d'utilité précédente.

Par la suite, nous formulons notre problème d'optimisation en considérant comme objectif la maximisation de la QoI et de l'équité de l'allocation des tâches et comme contraintes la consommation d'énergie ainsi que la redondance de données comme suit :

$$\begin{aligned}
 & \text{Maximize : } \lambda_1 \times U_{QoI}(X) + \lambda_2 \times J(X) \\
 & \text{subject to : } U_e(e_i - e_c \sum_{j=1}^M x_{ij}) \geq 0 \quad \forall i \in 1 \dots N \\
 & \text{and } \sum_{i=1}^N x_{ij} \leq 1 \quad \forall j \in 1 \dots M
 \end{aligned} \tag{1}$$

Où U_{QoI} est le produit des fonctions d'utilités de la précision temporelle et la complétude de données détaillées en [3], X est la matrice solution avec les lignes correspondant aux utilisateurs et les colonnes aux régions/ tâches de collecte et λ_1 et λ_2 sont les coefficient associés des fonctions objectifs. Notons que pour trouver la solution *Pareto*, ces coefficients doivent être positifs. Aussi, nous figeons par la première contrainte un niveau de batterie seuil, e_θ , en dessous duquel le dispositif n'est plus participant aux campagnes de collecte. Cela garantit aux utilisateurs l'énergie nécessaire pour l'utilisation quotidienne de leurs mobiles. Finalement, nous exigeons par la deuxième contrainte qu'une mesure précise est collectée par un seul utilisateur pour éviter la redondance des données.

3 F-QEMSS : Implémentation et Résultats

3.1 F-QEMSS : Algorithme basé sur la recherche taboue

Pour résoudre le problème d'optimisation (1), nous avons eu recours à l'algorithme de la recherche taboue destiné à la résolution des problèmes combinatoires. Cet algorithme dispose d'une liste taboue de solutions visitées, ce qui minimise le temps de calcul et accélère la convergence de notre modèle.

Nous modélisons la solution de notre algorithme par une matrice booléenne X ; où les lignes représentent les participants et les colonnes désignent les tâches à réaliser (régions de sensing). Ainsi, si $x_{i,j} = 1$, l'utilisateur i est désigné pour faire le sensing dans la région j . Aussi, nous formulons une matrice de référence X_{ref} pour vérifier la présence, $x_{ref}(i,j) = 1$, ou non, $x_{ref}(i,j) = 0$, de chaque participant dans une sous-région de sensing donnée. Enfin, nous calculons la QoI potentiellement acquise par un participant pour une tâche, $u_{i,j}(QoI)$. Les étapes décrites ci-dessous détaillent notre modèle, F-QEMSS :

- **Initialisation** : Pour générer une solution initiale, nous devons déterminer le participant p_i qui maximise la fonction objectif de (1) tout en vérifiant la contrainte énergétique, si valide, le participant est alors sélectionné et nous considérons la tâche suivante.
- **Formation du voisinage $N(X)$** : En se basant sur la solution initiale X_{init} , nous itérons les différentes combinaisons possibles d'allocation de tâches aux participants. Pour ce faire, nous utilisons la modification de la solution X actuelle pour en générer d'autres, définie comme mouvement M , et nous considérons dans ce travail la multiplication (XOR) : $X' = X \oplus M$, où M est une matrice de même taille que X avec tous les éléments nuls sauf ceux de l'ancienne et la nouvelle allocation.
- **Sélection de la solution** : Pour chaque solution générée dans $N(X)$, nous calculons la valeur de la fonction objectif, $\lambda_1 \times U_{QoI}(X) + \lambda_2 \times J(X)$, pour choisir la solution avec la valeur maximale.
- **Mise à jour de la liste taboue (TL)** : Nous ajoutons les attributs des solutions visitées à la TL pour éviter de retourner à ce voisinage de solutions.

Nous itérons les étapes précédentes jusqu'à l'affectation complète des tâches ou l'expiration du nombre d'itérations maximum de l'algorithme (TS). La solution générée est ainsi l'affectation équitable qui assure le maximum de QoI et l'équité aux participants au Crowd-Sensing.

3.2 Performances et résultats

3.2.1 Environnement de simulation

Nous articulons notre simulation autour d'une modélisation d'une zone de collecte urbaine de taille $4000m \times 6000m$. Pour la mobilité des utilisateurs, nous générons des traces de mobilité du Manhattan mobility model par le simulateur des traces, Bonnmotion en variant le nombre des participants entre 100 et 1000. En outre, l'énergie initiale des dispositifs est arbitrairement générée suivant une loi uniformément distribuée entre le seuil énergétique e_0 et 100%. Puisque nous considérons un compromis entre la QoI et l'équité, nous fixons les deux coefficients des fonctions objectifs à $\lambda_1 = \lambda_2 = 0,5$. Au total, nous avons effectué des simulations pendant 16 heures de sensing pour chaque nombre d'utilisateurs.

3.2.2 Performances

Pour évaluer notre modèle, nous nous référons à deux métriques : la valeur maximale de la qualité de données collectées et l'équité mesurée par l'indice de Jain. Nous comparons la performance de notre modèle F-QEMSS à celle de son prédécesseur non-équitable QEMSS [3], un modèle de sélection dynamique des participants (DPS) de l'état de l'art ainsi qu'à une méthode de sélection équitable qui affecte le même nombre de tâches de sensing à tous les participants sans considération de QoI défini par Equal Sensing (ES). Les résultats des simulations sont illustrés sur la Figure 2.

La Figure 2.a illustre la valeur moyenne du taux maximum de qualité de données atteint par les différents modèles d'attribution des tâches de collecte participative : DPS, F-QEMSS, QEMSS et ES. Naturellement, ce taux augmente en fonction du nombre d'utilisateurs présents dans la région. Notons que ES réalise le taux le plus bas de QoI, vu qu'il vise essentiellement à maximiser l'équité de l'affectation des tâches. DPS atteint une valeur moins importante que celle réalisée par QEMSS et F-QEMSS. Cela est dû à la stratégie de sélection gloutonne utilisée dans ce modèle, où l'on sélectionne non pas les meilleurs utilisateurs mais plutôt les premiers utilisateurs ayant une bonne qualité de contribution.

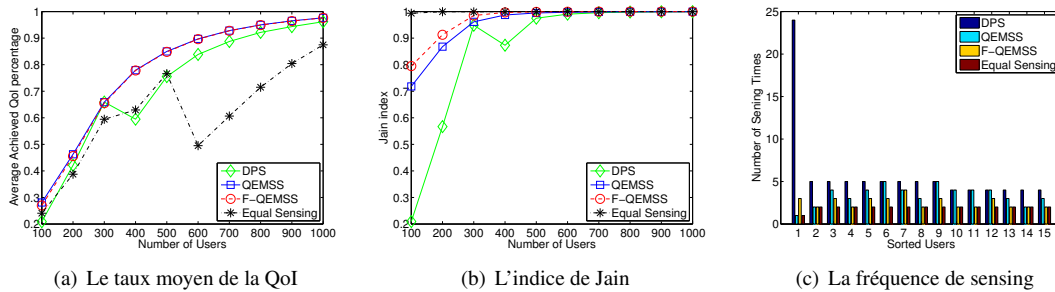


FIGURE 2: Évaluation de performance des différents modèles de déploiement en fonction du nombre de participants.

La Figure 2.b illustre l'évaluation de l'équité des différentes méthodes d'allocation de tâches, mesurée par l'indice de Jain. ES est la méthode la plus équitable vu qu'elle assigne le même nombre de tâches à tous les participants et elle atteint la valeur maximale de l'indice de Jain, à savoir 1. En revanche, la méthode la moins juste est l'algorithme DPS. Particulièrement, dans les zones de sensing à faible densité, plus de 50% des utilisateurs sont non satisfaits. Pour les deux autres modèles de déploiement, QEMSS et F-QEMSS, les valeurs d'indice de Jain varient entre 0,78 et 1. Par conséquent, la majorité des participants sont satisfaits en dédiant leurs ressources pour réaliser le maximum de QoI.

Pour conclure, il est clair que le modèle d'allocation des tâches de sensing, F-QEMSS, réalise un important taux de qualité de données et d'équité. Cela est aussi illustré par le graphique 2.c, qui décrit la distribution du nombre de tâches de sensing entre les 15 participants qui ont effectué le plus de campagnes parmi 500. Par définition, la fréquence de collecte est distribuée équitablement par la méthode Equal Sensing. Par ailleurs, les utilisateurs sélectionnés par l'algorithme DPS sont les moins satisfaits. Il est aussi à noter que F-QEMSS est plus équitable que QEMSS tout en préservant la même qualité de données.

4 Conclusion

Dans ce papier, nous avons présenté un modèle d'allocation équitable des tâches de collecte participative, F-QEMSS, pour répondre aux exigences de qualité des données chez les usagers sans abuser des ressources énergétiques des participants. Nous avons formulé le problème d'optimisation multi-objectif correspondant puis nous avons conçu une solution basée sur l'heuristique de la recherche taboue. Les résultats de simulation ont montré le fait que notre algorithme, comparé à d'autres modèles de déploiement de collecte participative, réalise un niveau important d'équité. En outre, F-QEMSS maintient le même niveau de QoI qu'atteignent les méthodes non-équitables. Par conséquent, nous réalisons un excellent compromis de satisfaction des demandeurs et des participants lors du processus de Crowdsensing. Nous allons étendre notre travail pour considérer une mobilité réaliste des utilisateurs et appuyer la précision du modèle proposé.

Références

- [1] N. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. Campbell, "A survey of mobile phone sensing," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 140–150, 2010.
- [2] X. Sheng, J. Tang, X. Xiao, and G. Xue, "Leveraging GPS-less sensing scheduling for green mobile crowd sensing," *IEEE Internet of Things Journal*, vol. 1, no. 4, pp. 328–336, 2014.
- [3] R. B. Messaoud and Y. Ghamri-Doudane, "QEMSS : A selection scheme for participatory sensing tasks," in *Protocol Engineering (ICPE) and International Conference on New Technologies of Distributed Systems (NTDS), 2015 International Conference on*, 2015, pp. 1–6.
- [4] T. Luo and C.-K. Tham, "Fairness and social welfare in incentivizing participatory sensing," in *IEEE Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, 2012, pp. 425–433.
- [5] Q.-T. Nguyen-Vuong, Y. Ghamri-Doudane, and N. Agoulmine, "On utility models for access network selection in wireless heterogeneous networks," in *NOMS*, 2008, pp. 144–151.