

Clustering par Budget: Algorithme et Tests

Fehmi Ben Abdesslem¹, Artur Ziviani², Marcelo Dias de Amorim¹, Petia Todorova³

¹UPMC Univ Paris 06, France ² LNCC, Brésil ³Institut FOKUS, Allemagne

{fehmi,amorim}@rp.lip6.fr ziviani@lncc.br petia.todorova@fokus.fraunhofer.de

Abstract— Former des groupes de nœuds (clusters) de manière distribuée dans un réseau sans fil spontané est essentiel au passage à l'échelle de nombreux protocoles de communication. Afin d'obtenir une taille moyenne de clusters prédéfinie, nous proposons un algorithme basé sur la distribution de jetons en fonction du degré de connectivité des voisins. Nous montrons à l'aide de simulations que les clusters ainsi formés présentent une taille plus proche du budget initial qu'avec les propositions antérieures.

I. INTRODUCTION

Comme d'autres algorithmes de formation de clusters basés sur des budgets [1], notre algorithme commence par l'affectation de β jetons (budget) à un ensemble de nœuds initiateurs, aléatoirement choisis, qui feront également office de *cluster-heads*. Le budget β correspond à la taille idéale des clusters. Quant au choix des nœuds initiateurs, il se fait de manière distribuée en armant une minuterie sur chaque nœud, réglée avec un temps aléatoire. A l'expiration de sa minuterie, un nœud déclenche la formation du cluster en distribuant le budget initial à ses voisins directs. Lorsqu'un nœud reçoit un jeton avant l'expiration de sa minuterie, celle-ci est aussitôt désarmée. Les voisins vont à leur tour recruter de nouveaux nœuds pour former le cluster, en redistribuant à leurs voisins respectifs les jetons qui leur ont été attribués, et ceci récursivement jusqu'à épuisement du budget. La convergence de l'algorithme est garantie par la consommation d'un jeton par chacun des nœuds recrutés. Dans le meilleur des cas, chaque cluster ainsi formé a une taille égale à β .

Former des clusters dont la taille est la plus proche possible du budget de départ est une problématique clé. Des algorithmes de clustering existants basés sur des budgets tels que Rapid ou Persistent [1], distribuent aveuglément le budget au voisinage, à parts égales. Ainsi, un nœud n'ayant aucun autre voisin à recruter peut se voir allouer inutilement des jetons à redistribuer. Ces jetons perdus auraient pu être utilisés par un autre nœud avec un fort degré de connectivité mais n'ayant pas reçu suffisamment de jetons pour étendre le cluster à ses voisins. Ce cas de figure apparaît souvent dans des topologies dans lesquelles la densité varie à travers le réseau.

Nous proposons un nouvel algorithme de clustering basé sur des budgets, qui prend en compte les degrés de connectivité de chaque nœud avoisinant pour pondérer la distribution des jetons. Cet algorithme, que nous appelons PBC (*Potential-Based Clustering*) [2] [3], produit des clusters dont la taille est bien plus proche du budget β par rapport aux approches existantes.

II. CLUSTERING BASÉ SUR LA CONNECTIVITÉ (PBC)

Dans les algorithmes de clustering basés sur des budgets, un nœud initiateur dispose au préalable d'un budget alloué

β , qu'il commence par consommer en s'attribuant une unité, avant de distribuer les $\beta - 1$ restants à ses voisins. Ses voisins consomment à leur tour une unité du budget reçu, et redistribuent les jetons restants à leurs voisins (en dehors de la source). Les redistributions sont ensuite effectuées récursivement jusqu'à épuisement du budget ou absence de voisins disponibles.

La distribution récursive du budget à parts égales provoque inévitablement une affectation excessive de jetons à certains voisins, entraînant la perte de jetons et la réduction de la taille du cluster en formation. Se limiter à un simple retour de ces jetons superflus, comme dans l'algorithme Persistent de Krishnan et al., permet d'augmenter sensiblement la taille moyenne des clusters, mais provoque la formation de clusters étendus et enchevêtrés les uns aux autres. Ce phénomène s'explique par une mauvaise distribution des jetons, qui entraîne l'isolement de nœuds en marge des clusters. Ces nœuds isolés, dont la plupart des voisins ont déjà été recrutés par la formation d'un précédent cluster, ne trouveront pas assez de voisins disponibles pour écouler l'ensemble du budget. Afin d'éviter la perte de budget et la création de clusters de taille réduite, nous proposons, en plus de retourner les jetons superflus, de ne plus les distribuer aveuglément.

Un pré-requis à l'algorithme est donc la connaissance par chaque nœud du degré de connectivité de ses voisins. Celui-ci peut être facilement déterminé en utilisant les balises envoyées régulièrement par les nœuds. Soit $\mathbf{N}(X) = \{x_1, x_2, \dots\}$ l'ensemble des voisins de X . La taille de cet ensemble définit le potentiel du nœud X , noté $\pi(X) = |\mathbf{N}(X)|$.

Soit A un nœud disposant d'un budget β_A . La quantité de jetons alloués à son voisin $a_i \in \mathbf{N}(A)$ est calculée de la manière suivante :

$$\beta_{a_i} = \left\lfloor \frac{(\beta_A - 1)\pi(a_i)}{\sum_{j=1}^{|\mathbf{N}(A)|} \pi(a_j)} \right\rfloor. \quad (1)$$

Afin de former des clusters plus compacts, et éviter ainsi l'apparition de clusters de petite taille créés par des nœuds isolés, notre algorithme inclut les mécanismes suivants:

- Les jetons non distribués par l'équation 1 sont répartis uniformément, en commençant par les voisins de faible potentiel.
- Un jeton est distribué à chaque voisin avant d'appliquer l'équation 1.
- Lorsque le dernier voisin disponible d'un nœud est recruté par un cluster, ce nœud s'associe également au cluster en formation.

Nous montrons dans la figure 1 comment PBC se comporte comparé à Rapid et Persistent, avec un budget $\beta = 13$. Pour

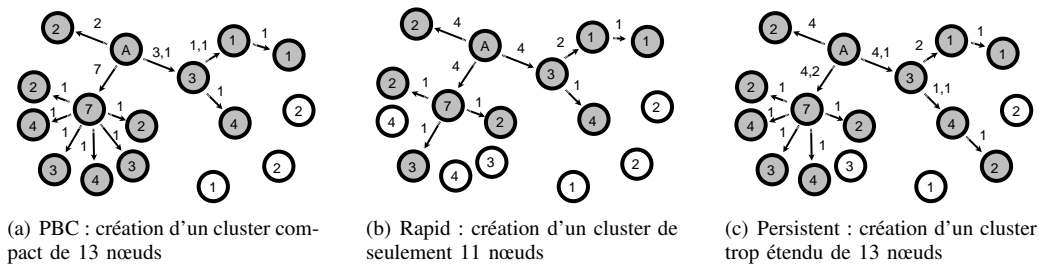


Fig. 1. Comparaison illustrative des différents algorithmes. L'inondation est initiée par le nœud A. Les liens à plusieurs valeurs sont sujets à de multiples distributions consécutives et l'étiquette de chaque nœud correspond à son potentiel.

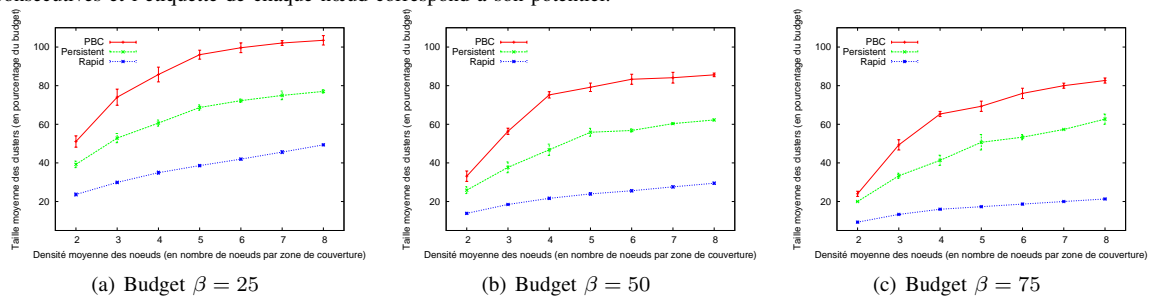


Fig. 2. Taille moyenne des clusters en fonction de la densité des topologies aléatoires.

plus de clarté, seule la distribution issue de la formule 1 a été utilisée, sans les mesures complémentaires destinées à éviter que des nœuds adjacents ne créent des clusters de petite taille. La figure montre que notre algorithme obtient le même résultat maximal que Persistent, mais s'efforce d'attribuer des jetons proportionnellement au degré de connectivité de chaque voisin afin de former des clusters plus compacts. Ainsi, les 3 nœuds encore disponibles sont moins dispersés dans la topologie, et peuvent donc appartenir à un même cluster. En revanche, Persistent sera plus souvent confronté à des nœuds isolés, qui devront chacun de leur côté créer un nouveau cluster en initiant une autre inondation, diminuant ainsi la taille moyenne des clusters.

III. SIMULATIONS

Pour comparer les algorithmes, des simulations ont été effectuées sur des topologies générées aléatoirement. Le simulateur utilisé est NS-2, et les topologies aléatoires sont générées selon une distribution uniforme, sur une surface circulaire de 2500m de rayon, chaque nœud ayant une couverture radio de 100m. La densité moyenne de la topologie varie de 2 à 8 nœuds par zone de couverture, soit de 1250 à 5000 nœuds. Pour chaque densité, une nouvelle topologie aléatoire est générée. Les trois algorithmes sont ensuite exécutés sur cette même topologie.

L'efficacité d'un algorithme de clustering basé sur les budgets s'évalue en fonction de la taille des clusters formés. Le cas idéal serait d'obtenir des clusters de taille égale au budget. Cependant, la propagation des jetons ne permettant pas toujours d'atteindre une telle taille pour tous les clusters, la taille moyenne des clusters reste en général bien inférieure au budget. Ainsi, plus cette taille moyenne est proche du budget, plus l'algorithme est efficace. Les résultats des simulations montrent que la taille moyenne des clusters est jusqu'à 20% plus importante lorsque les jetons sont distribués en tenant compte des potentiels, pour une densité de 8 nœuds par zone de couverture, et un budget de 75. Cependant, pour la même

densité, l'amélioration est moins nette pour des budgets plus faibles: 9% pour un budget de 50, et aucune amélioration significative pour un budget de 25. Pour obtenir des résultats quel que soit le budget, notre algorithme comprend également des mécanismes simples, décrits dans la partie précédente (e.g. flexibilité du budget). La figure 2 représente les résultats de notre algorithme avec l'ensemble de ses mécanismes, et pour un budget fixé à $\beta = 25, 50$ et 75 . Chaque valeur correspond à une moyenne sur 5 simulations. L'algorithme Rapid est le moins efficace puisqu'il distribue les jetons aveuglément sans mécanisme de retour. En renvoyant les jetons non alloués, Persistent parvient à de meilleurs résultats. Cependant, la taille moyenne des clusters est de 33.8 % plus importante en utilisant notre algorithme PBC (pour un budget de 75 et une densité de 8), mais également 35.48 % et 33.6 % pour des budgets plus faibles de 50 et 25, respectivement.

IV. CONCLUSION

Les algorithmes de clustering basés sur des budgets peuvent être optimisés en prenant en compte les degrés de connectivité des voisins lors de la distribution des jetons. Pour cela, nous proposons l'algorithme PBC (*Potential-Based Clustering*), qui comprend également d'autres mécanismes permettant d'obtenir des clusters de taille plus proches du budget. Des simulations confirment l'efficacité de cette approche sur des topologies uniformément aléatoires. Les travaux en cours visent à définir de manière distribuée le meilleur sous-ensemble de nœuds initiateurs optimisant les performances.

REFERENCES

- [1] R. Krishnan and D. Starobinski, "Efficient clustering algorithms for self-organizing wireless sensor networks," *Ad Hoc Networks*, vol. 4, no. 1, pp. 36–59, 2006.
- [2] F. Ben Abdesslem, A. Ziviani, M. Dias de Amorim, and P. Todorova, "Looking around first: Localized potential-based clustering in spontaneous networks," *IEEE Communications Letters*, vol. 11, no. 8, pp. 653–655, 2007.
- [3] F. Ben Abdesslem, A. Ziviani, M. Amorim, and P. Todorova, "Clustering pour réseaux spontanés basé sur le degré de connectivité," in *ALGO-TEL'08*, Saint-Malo, France, May 2008.