
Modélisation et Simulation des Systèmes

Tutoriel 1:
Simulation de K-moyennes

Dr. Manh Hung Nguyen

¹Posts and Telecommunications Institute of Technology (PTIT), Hanoi, Vietnam

²UMI UMMISCO 209 (IRD/UPMC), Hanoi, Vietnam

Algorithme de K-moyennes

Étant donné un ensemble de points (x_1, x_2, \dots, x_n) , on cherche à partitionner les n points en k ensembles $S = \{S_1, S_2, \dots, S_k\}$ ($k \leq n$) en minimisant la distance entre les points à l'intérieur de chaque partition :

$$\arg \min_{\mathbf{S}} \sum_{i=1}^k \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \boldsymbol{\mu}_i\|^2$$

où μ_i est la moyenne des points dans S_i .

Algorithme de K-moyennes (2)

- Choisir k points qui représentent la position moyenne des partitions $m_1(1), \dots, m_k(1)$ initiales (au hasard par exemple) ;
- Répéter jusqu'à ce qu'il y ait convergence :
 - assigner chaque observation à la partition la plus proche :

$$S_i^{(t)} = \left\{ \mathbf{x}_j : \|\mathbf{x}_j - \mathbf{m}_i^{(t)}\| \leq \|\mathbf{x}_j - \mathbf{m}_{i^*}^{(t)}\| \forall i^* = 1, \dots, k \right\},$$

- mettre à jour la moyenne de chaque cluster :

$$\mathbf{m}_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{\mathbf{x}_j \in S_i^{(t)}} \mathbf{x}_j.$$

Algorithme de K-moyennes (3)

- *Comment peut-on simuler la fonctionnement de cette algorithme?*

Problématique

- *Étant donné une surface représentant l'espace du problème*
- *Un ensemble des points sont représentées par un ensemble des points sur cette surface*
- *Choisir K couleurs pour les K clusters*
- *À chaque étape: changer la couleur des points à base sur leur distance aux K moyennes*
- *Arreter quand il n'y a plus de changement du couleur dans un étape*



Modélisation



Extraire des agents

Combient d'agent possible dans le système?

- *Les N points*
- *Les K moyennes*
- *Autres: surface, système?*

Agent: Point

Attributes:

- *Les coordonnées (x,y)*
- *La couleur actuelle (de quelle moyenne)*

Activités:

- *Recevoir les coordonnées de toutes les moyennes*
- *Calculer la distance à chaque moyenne pour choisir la moyenne la plus proche*
- *Changer la couleur à base de la moyenne la plus proche*
- *Renvoyer leur coordonnées à la moyenne choisie*
- *Communiquer avec autre point? - non*
- *Communiquer avec les moyennes? - oui*

Agent: Moyenne

Attributes:

- *Les coordonnées (x,y)*
- *La couleur actuelle*

Activités:

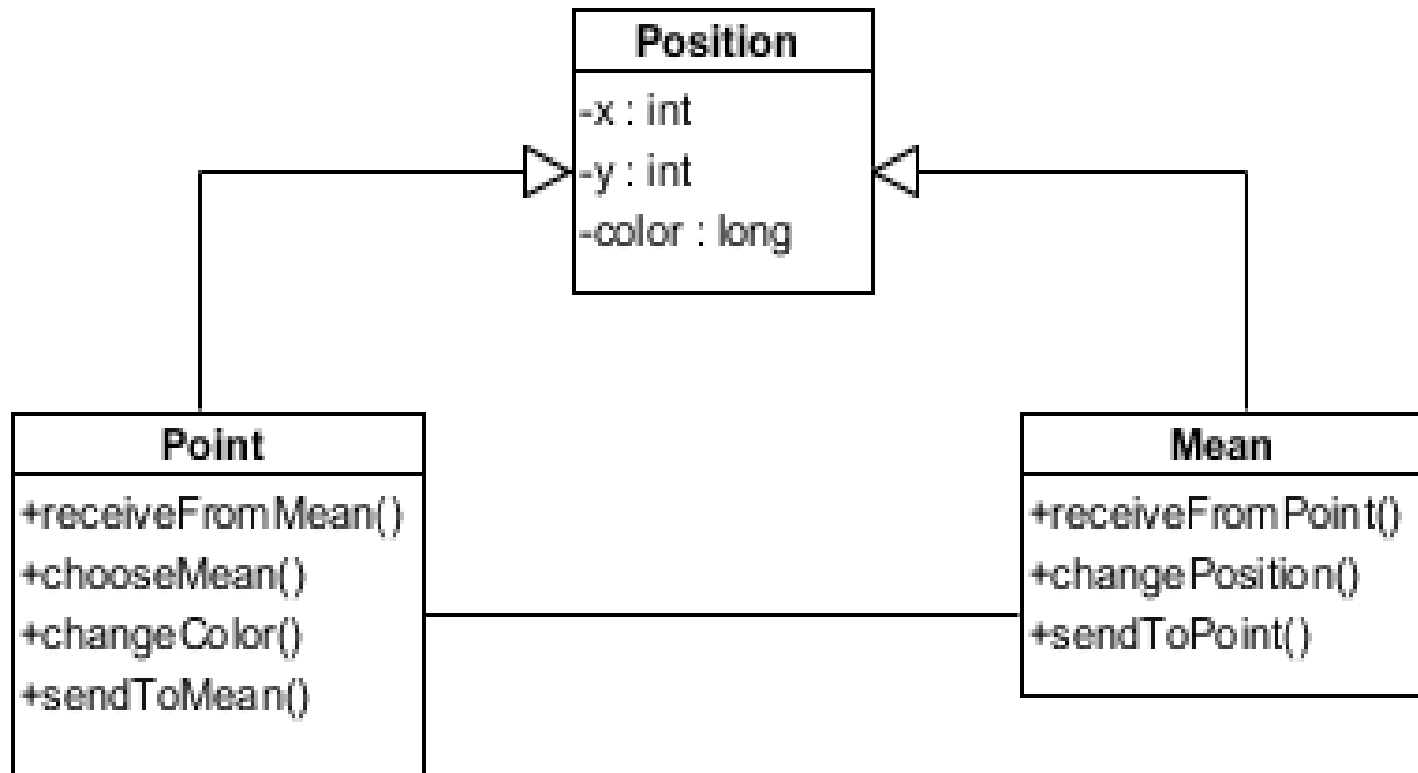
- *Envoyer leur coordonnées à toutes les points*
- *Recevoir les coordonnées des points de meme couleur*
- *Calculer les nouvelles coordonnées après l'étape*
- *Changer les coordonnées après l'étape*
- *Communiquer avec des points? - oui*
- *Communiquer avec autre moyennes? - non*

Agent: Moyenne (2)

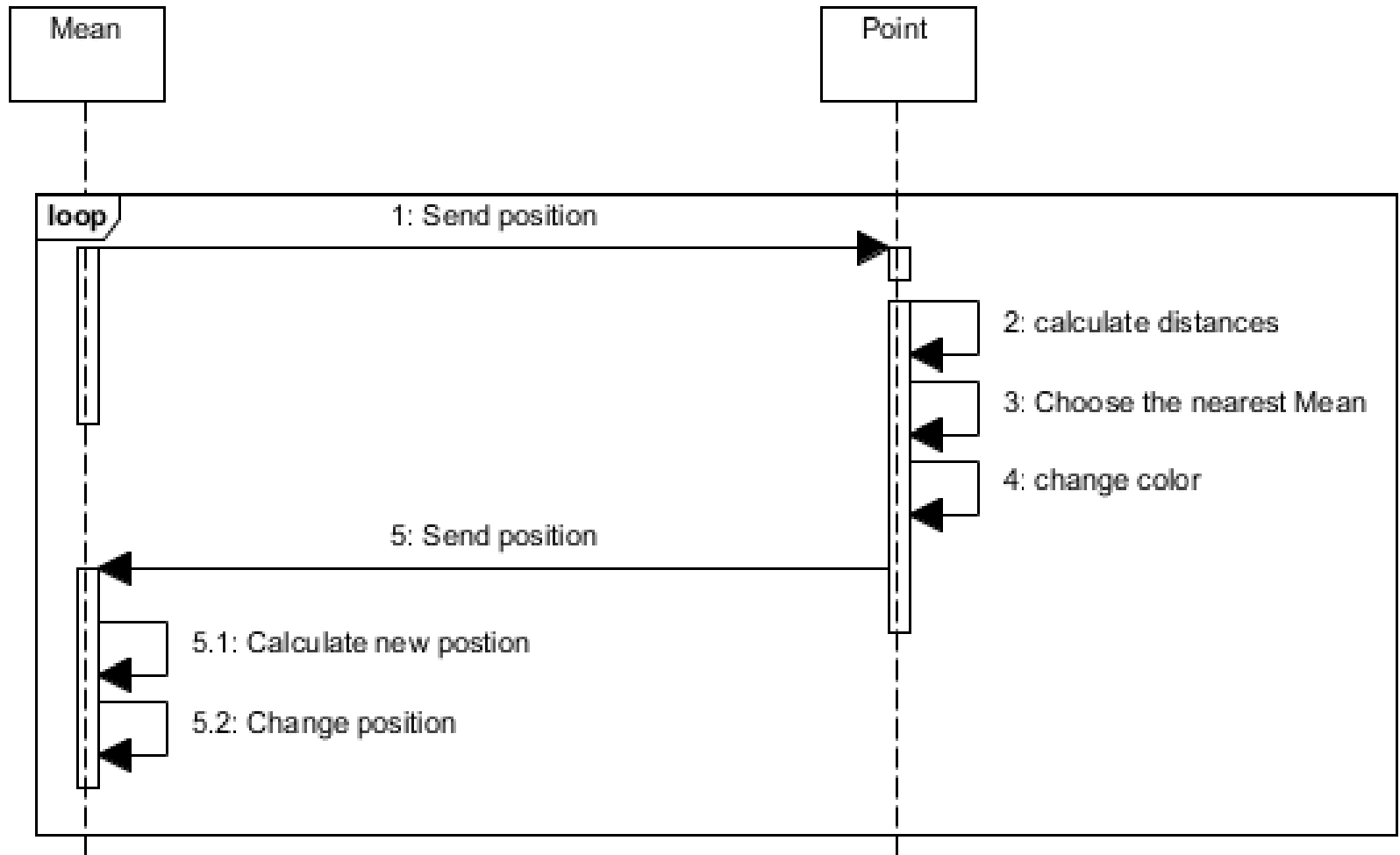
Comment peut-elle changer les coordonnées après l'étape?

- *Soit par se déplacer vers la nouvelle position*
- *Soit par se mortir celle à l'ancienne position, puis créer une nouvelle à la nouvelle position*
- *Quelle est la meilleure solution?*

Les agents



Les activités dans une étape





Simulation



Scénario (1)

Initial:

- *Créer les N points au hasard sur la surface*
- *Choisir les K moyennes au hasard à partir de N points*
- *Définir une couleur pour une moyenne*

Scénario (2)

Repéter:

- *Les moyennes envoient leur position aux points*
- *Les points calculent leur distance à chaque moyenne, et puis, choisissent la plus proche*
- *Les points changent leur couleur, et puis envoient leur position à la moyenne la plus proche*
- *Les moyennes reçoivent les positions des points proches et calculent la nouvelle position, et puis y se déplacent*

Condition d'arrêt:

- *Il n'y a plus de changement de position des moyennes*

Dans GAMA

Définir les agents:

- *Les attributes*
- *Les actions*

Réaliser les interaction entre les agents:

- *L'envoi de position de la moyenne aux points*
- *L'envoi de position du point à la moyenne la plus proche*

Lancer la simulation

Définir:

- *La condition d'arrêt*
- *Les paramètres sorties?*