

---

Génie Industriel et Mathématiques Appliquées

---

---

# Projet de Modélisation Stochastique Appliquée

---

SPEISMANN Matthieu et POURMY Thomas

Ecole Nationale Supérieure des Mines de Nancy

Année scolaire 2025-2026

# I Modélisation du problème

- Les colonies sont les lieux où un vélo peut se trouver à un instant  $t$  donné. Ainsi, les colonies sont ici :
  - Les stations de vélo ( $S_i$ ) $_{i \in [1, N]}$
  - Les trajets orientés ( $T_{i,j}$ ) $_{i,j \in [1, N]^2, i \neq j}$  entre les stations.

Cela permet de définir l'état du système à l'instant  $t$  par :

$$X(t) = (X_{S_1}(t), \dots, X_{S_N}(t), X_{T_{1,2}}(t), X_{T_{1,3}}(t), \dots, X_{T_{N,N-1}}(t))$$

où, à l'instant  $t$ ,  $X_{S_i}(t)$  le nombre de vélo à la station  $S_i$  et  $X_{T_{i,j}}(t)$  le nombre de vélo sur le trajet orienté  $T_{i,j}$ .

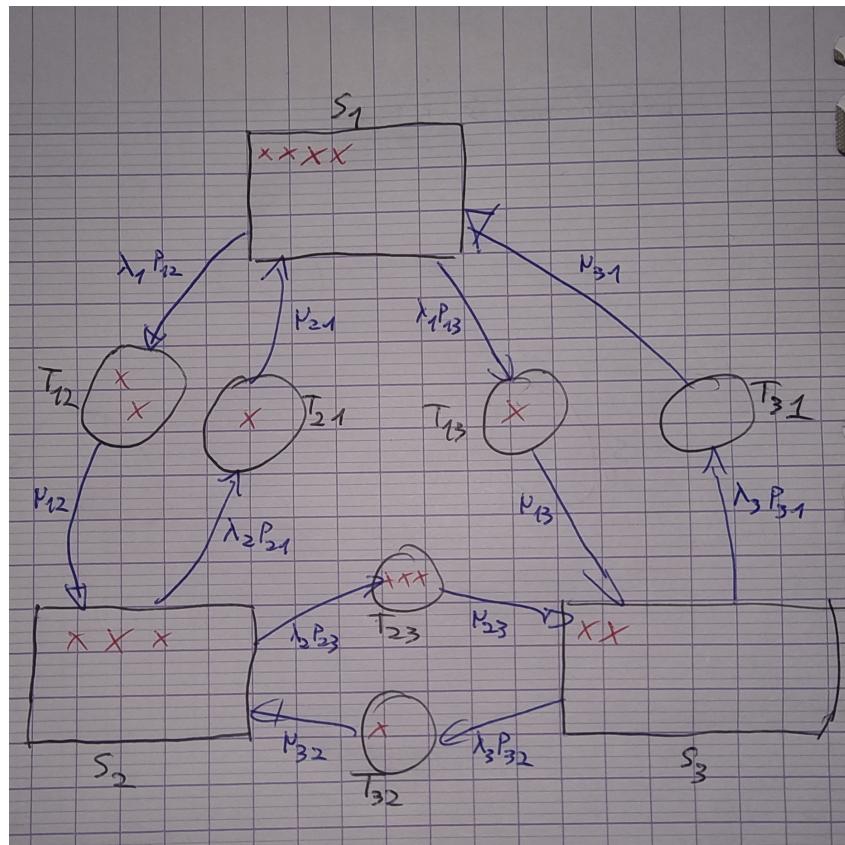
De plus, le nombre de vélo total dans le système est constant et égal à  $K$  d'où la contrainte

$$\sum_{i=1}^N X_{S_i}(t) + \sum_{i \neq j} X_{T_{i,j}}(t) = K$$

L'espace d'état discret fini utilisé est donc

$$\mathbb{E} = \left\{ x = (X_{S_1}, \dots, X_{S_N}, X_{T_{1,2}}, \dots, X_{T_{N,N-1}}) \in \mathbb{N}^{N+N(N-1)} \mid \sum_i X_{S_i} + \sum_{i \neq j} X_{T_{i,j}} = K \right\}$$

- Voici un schéma d'exemple avec  $N = 3$  stations, et donc 6 trajets orientés possibles, qui forment ensemble 9 colonies. De plus, on a ici  $K = 17$  vélos.



3. On a besoin de définir des vecteurs de transition  $e_{S_i}$  et  $e_{T_{i,j}}$  pour les départs de station et les trajets tel que :

$\forall x = (X_{S_1}, \dots, X_{S_N}, X_{T_{1,2}}, \dots, X_{T_{N,N-1}})$  un état du système.

- $x \pm e_{S_i} = (X_{S_1}, \dots, X_{S_i} \pm 1, \dots, X_{S_N}, X_{T_{1,2}}, \dots, X_{T_{N,N-1}})$
- $x \pm e_{T_{i,j}} = (X_{S_1}, \dots, X_{S_N}, X_{T_{1,2}}, \dots, X_{T_{i,j}} \pm 1, \dots, X_{T_{N,N-1}})$

On remarque qu'il existe deux types de transitions possibles :

- Un départ de station :

Si un vélo est disponible, un client commence un trajet depuis la station  $S_i$  avec un taux  $\lambda_i$ .

Ensuite, il a une probabilité  $p_{i,j}$  de se diriger vers la station  $S_j$  et ainsi de faire le trajet  $T_{i,j}$ .

Donc ce type de transition est défini dans le générateur par  $Q(x, x - e_{S_i} + e_{T_{i,j}}) = \lambda_i p_{i,j} 1_{x_{S_i} > 0}$ .

- Une fin de trajet :

Chaque trajet de  $i$  vers  $j$  met un temps exponentiel de paramètre  $\mu_{i,j}$  à se terminer.

Donc ce type de transition est défini dans le générateur par  $Q(x, x + e_{S_j} - e_{T_{i,j}}) = \mu_{i,j} x_{T_{i,j}}$ .

On conclut donc que le générateur infinitésimal  $Q$  du processus de Markov est défini par :

$$Q = \begin{cases} Q(x, x - e_{S_i} + e_{T_{i,j}}) = \lambda_i p_{i,j} 1_{x_{S_i} > 0} \\ Q(x, x + e_{S_j} - e_{T_{i,j}}) = \mu_{i,j} x_{T_{i,j}} \\ Q(x, x) = - \sum_{y \neq x} Q(x, y) \end{cases}$$

4. Les paramètres du modèle sont les suivants :

- $\lambda_i$  sont les taux d'arrivée de client à la station  $S_i$ . En clients par heure (équivalent à intensité de départ de vélo par station)
- $\mu_{i,j}$  sont les taux d'arrivée du trajet de  $i$  vers  $j$ , inverses du temps moyen du trajet de la station  $i$  vers la station  $j$ . En heures $^{-1}$ .
- $p_{i,j}$  les probabilités qu'un client qui part de la station  $i$  réalise un trajet vers la station  $j$ . Sans unité.
- $K$  le paramètre global du nombre total de vélo dans le système.

## II Calibration des paramètres

Avec les données fournies, on peut estimer les paramètres du modèle comme :

- $\lambda_i$  est estimé par le nombre de départs de vélo par heure à la station  $S_i$ .
- $\mu_{i,j}$  est estimé par l'inverse du temps moyen du trajet de la station  $i$  vers la station  $j$ .
- $p_{i,j}$  est estimé par la probabilité d'aller en  $j$  partant de  $i$ .
- $K$  est le nombre total de vélo dans le système, égal à la somme des vélos présents dans les stations à l'instant initial.

## III Simulation

1. Voir le code de la simulation dans le notebook disponible sur le dépôt :

<https://github.com/Matthieu-Speismann/Simulation-Velib-Stochastique>

2. En calculant d'une variable aléatoire qui vaut 1 si une station est vide après 150 heures, et 0 sinon, on peut calculer cette proportion. Avec 200 simulations, on obtient un résultat de 0.735

3. On calcule un intervalle de confiance à 95% pour la proportion de temps où une station est vide.

Avec  $p$  la proportion calculée à la question précédente et  $M$  le nombre de simulations, on trouve l'intervalle de confiance suivant :

$$\left[ p - 1.96 \sqrt{\frac{p(1-p)}{M}}, p + 1.96 \sqrt{\frac{p(1-p)}{M}} \right] = [0.674, 0.796]$$

4. En conservant l'historique de l'état des stations au cours d'une simulation, on détermine numériquement cette valeur, ici égale à 0.748.

5. En vertu de l'ergodicité du processus de Markov utilisé ainsi que par la propriété PASTA, on s'attend à ce que ces deux résultats coïncident et soient tout deux proches de la probabilité stationnaire d'une station vide.

Cela est vérifié ici car, en effet, les deux valeurs obtenues sont proches, en particulier, le pourcentage du temps où il existe une station vide est compris dans l'intervalle de confiance de la probabilité d'avoir une station vide après 150 heures.