Mean field analysis of an incentive algorithm for a closed stochastic network

Bianca Marin Moreno, Christine Fricker, Hanene Mohamed, Amaury Philippe, Martin Trepanier

MEAN FIELD ANALYSIS OF AN INCENTIVE ALGORITHM FOR A CLOSED STOCHASTIC NETWORK

BIANCA MARIN MORENO, CHRISTINE FRICKER, HANENE MOHAMED, AMAURY PHILIPPE, AND MARTIN TREPANIER

Abstract. The paper deals with a load-balancing algorithm for a closed stochastic network with two zones with different demands. The algorithm is motivated by an incentive algorithm for redistribution of cars in a large-scale car-sharing system. The service area is divided into two zones. When cars stay too much long in the low-demand zone, users are encouraged to pick up them and return them in the high-demand zone. The zones are divided in cells called stations. The cars are the network customers. The mean-field limit solution of an ODE gives the large scale distribution of the station state in both clusters for this incentive policy in a discrete Markovian framework. An equilibrium point of this ODE is characterized via the invariant measure of a random walk in the quarter-plane. The proportion of empty and saturated stations measures how the system is balanced. Numerical experiments illustrate the impact of the incentive policy. Our study shows that the incentive policy helps when the high-demand zone observes a lack of cars but a saturation must be prevented especially when the high-demand zone is small.

Keywords. Large scale analysis, mean-field, car-sharing, incentive algorithm, stochastic network, cluster, load balancing.

Motivation. Car-sharing, a practice that is gaining ground in urban areas, comes to meet ecological, economic and practical imperatives. For a decade it has been becoming an alternative mode of transportation. The principle is that a given number of vehicles made available to users at stations or in public space in a given geographical area to make trips. The user picks up a vehicle if available, makes his trip and then drops it off at his destination.

For the operator, managing such systems is far from simple. The randomness due to the user arrivals as well as to the trips generates an imbalance in the system: Some areas are more or less served by vehicles throughout the day, depending on whether they are residential areas, on hill or in the city center for example. Thus, the users may find themselves without an available vehicle, which alters the efficiency of the system. Rebalancing the network by better distributing the vehicles, in other words, bringing them back where it is needed, is a major issue for operators. The usual techniques are either active, such as using trucks to move bikes or drivers for cars, or passive, such as incitative policies that encourage users to move vehicles themselves on their trips. We can cite the example of Velib+ which offered extra time for returning bikes in uphill stations of the Parisian bike-sharing system or the bikes Angel's Rewards program developed in NYC allowing to earn free day passes and membership extensions.

Gift incentive policy. This paper deals with an incentive policy implemented by Communauto on its free floating car-sharing system in Montreal. In the geographical area, a small zone is identified as a high-demand zone by the operator. Moreover, some cars remain stationary for too long in the rest of the service area with low-demand called normal zone while users cannot find available cars in the high-demand zone. In order to bring back these stagnant cars from the normal zone to the high-demand zone, Communauto designates them as *gifts* on its application and offers 30 free minutes on the trip if the user returns the *gift* to the high-demand zone. This policy is called here *gift policy.*

Aim of the paper. The aim is to study the impact of the incentive policy implementing a trip discount to move some cars to a high-demand area. For that, a probabilistic model is proposed for such a system as a large closed stochastic network of interacting particles which are cars and gifts. The service area is divided into cells, called here *stations*, which are nodes of the network, plus extra-nodes containing moving cars and gifts.

Results. We investigate in a Markovian framework the steady state of these stations. Although it exists an invariant measure for this irreducible Markov process on a finite state space for a fixed number of stations, it remains untractable. The idea is to deal with the approximation as the number of stations and cars get large together, called mean-field limit. Indeed, the states of the stations are asymptotically independent and their common distribution is given as a solution of an ODE. See Proposition 2.1. The equilibrium point of the ODE gives the long-time limit. The special case of a model without incentive policy corresponds to the two-cluster model studied in [**?**] where the equilibrium point is unique and well determined. Here is a practical application of this framework. For the *gift policy*, Proposition 2.2 gives a characterization of the equilibrium point as a function of the invariant measure of a random walk in the quarter-plane. It is a first step to address the problem of existence and uniqueness of the equilibrium point.

Performance. Our performance criterion is to minimize the proportion of empty or saturated stations, called for short *problematic*, in order to maximize the efficiency of the system. Since no closed-form solution for the previous invariant measure is derived, we perform a numerical solution of a multidimensional equation for the system with incentive policy. We compare it with the analytical solution of the model without incentives. We study the impact of the policy in the case where everyone follows incentives. This impact is significant when the high-demand zone lacks cars. The risk is to overload it, especially if it is small.

1. The model

1.1. **Model description.** In this following description and in the whole paper, a car is always a *normal car* and not a gift. We propose a simplified stochastic two-cluster model for car-sharing when including the gift policy. The principle of the model is the following.

- A user arrives in a station of cluster i with rate λ_i, where $i \in \{1, 2\}$. As the rate of user arrivals is larger in cluster 1 than cluster 2, $\lambda_1 > \lambda_2$.
- If the user arrives at a station in cluster 1 where there is an available car, the user picks it up to a trip. Otherwise he/she leaves the system.
- Every car parked in cluster 2 becomes a *gift* after a random time with mean $1/\delta$.
- When a user arrives in a station of cluster 2, if there is an available *gift* and an available car in this station, he/she picks up a gift with probability p, and a car with probability $1 - p$. If there is just one of the resources (gift or car), the user picks it up. Otherwise he/she leaves the system.
- When a car trip ends (at rate μ), the user chooses cluster i with probability c_i, then he/she chooses a station at random in this cluster to park the car.
- When a *gift* trip ends (at rate μ_c), the user returns the *gift* car to any station in cluster i with probability q_i. The *gift* parked appears then as a car on the app.
- A station in cluster i has capacity K_i. If the station chosen is full, the user makes another trip until finding a station with an available parking space.

In our model, the inter-arrival times of users, trip times and times to become a gift are all independent with exponential distribution. See Figure 1 for an illustration of the model.

1.2. **Notations.** Let us summarize the notations. For all the following, $i \in \{1, 2\}$ is the cluster type.

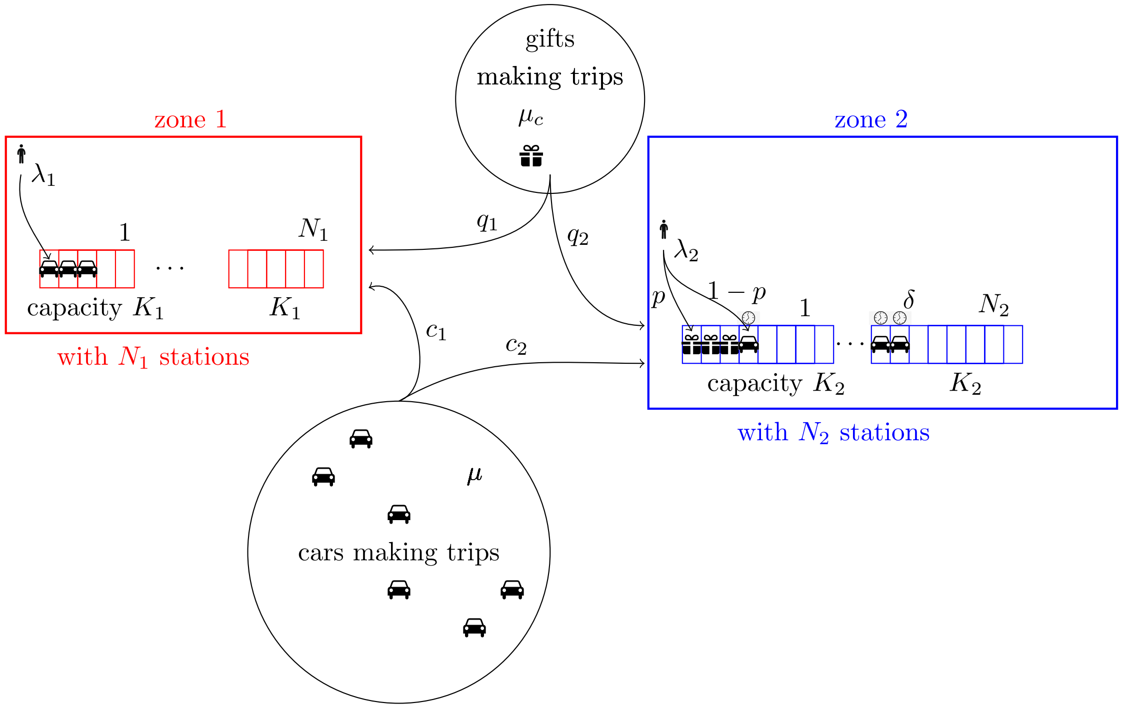


FIGURE 1. Illustration of the model with gifts.

- N_i is the number of stations in cluster i.
- $N = \sum_i N_i$ is the total number of stations.
- $\alpha_i = \lim_{N\to+\infty} N_i/N$ is the limiting proportion of stations in cluster i.
- K_i is the capacity of a station in cluster i.
- M is the total number of cars.
- $s = \lim_{N\to+\infty} M/N$ is the limiting mean number of cars per station, called *fleet size parameter*.
- λ_i is the rate of user arrivals at a station in cluster i.
- $1/\mu$ is the mean trip time for a normal car.
- $1/\mu_c$ is the mean trip time for a *gift*.
- δ is the rate at which a car in a station of cluster 2 becomes a *gift*.
- p is the probability that a user takes a gift when cars and gifts are both available.
- q_i is the probability that a user returns the gift in cluster i.
- c_i is the probability that a user returns his normal car in cluster i.

1.3. **Queueing formulation.** The system can be described as a closed stochastic network. The nodes of the network are a set of $N = N_1 + N_2$ one-server queues of finite capacity (the stations), divided in two clusters, cluster 1 (zone with high demand) with N_1 stations of capacity K_1, cluster 2 (normal zone) with N_2 stations of capacity K_2, plus two infinite-server queues, i.e. the two nodes containing respectively cars and gifts making a trip. The service times at the queues have exponential distribution with parameters respectively λ_1, λ_1, μ and μ_c. According to the queueing vocabulary, there are M customers of two classes: cars and gifts, and a routing matrix given by the previous description.

However, this is not a Jackson network because there are additional transitions since a car in a station of cluster 2 becomes a gift at rate δ and a gift arriving at a station from the infinite-server node becomes a car. It does not fit in this classical framework because of these changes of customers classes. Note that in the case without incentive policy ($\delta \to 0$), the model is a Jackson network since there are no gifts. Such a model is known as a two-cluster bike-sharing system studied in [**?**] and [**?**]. Section 2.1.1 is devoted to this case called *model without gifts*.

1.4. **The Markov process.** The state process is

$$\big(X_{1,n}(t), X_{2,m}(t), C_m(t), Z^N(t),\ 1 \le n \le N_1 \text{ and } 1 \le m \le N_2\big)$$

where

- $X_{1,n}(t)$ is the number of cars at a station n in cluster 1 at time t,
- $X_{2,m}(t)$ is the number of cars at a station m in cluster 2 at time t,
- $C_m(t)$ is the number of *gifts* at a station m (necessarily in cluster 2) at time t and
- $Z^N(t)$ is the number of *gifts* making a trip at time t.

Note that the number of cars making a trip at time t is equal to

$$M - \sum_{n=1}^{N_1} X_{1,n}(t) - \sum_{m=1}^{N_2} X_{2,m}(t) - Z^N(t).$$

As we deal with a two-cluster model, it is sufficient to study the behavior of one station in each cluster. It amounts dealing with the empirical measure process

$$(Y^N(t)) = \left(Y_{1,j}^{N_1}(t), Y_{2,k,l}^{N_2}(t), \frac{Z^N(t)}{N}, j \in \chi_1, (k,l) \in \chi_2\right)$$

where $Y_{1,j}^{N_1}(t)$ is the proportion of stations with j cars in cluster 1 and $Y_{2,k,l}^{N_2}(t)$ is the proportion of stations with k cars and l *gifts* in cluster 2, defined by

$$Y_{1,j}^{N_1}(t) = \frac{1}{N_1}\sum_{n=1}^{N_1} \mathbf{1}_{\left\{X_{1,n}(t)\right)=j\}} \text{ and } Y_{2,k,l}^{N_2}(t) = \frac{1}{N_2}\sum_{m=1}^{N_2} \mathbf{1}_{\left\{\left(X_{2,m}(t), C_m(t)\right)=(k,l)\right\}}$$

where $\chi_1 = \{j \in \mathbb{N},\ j \le K_1\}$ and $\chi_2 = \{(k,l) \in \mathbb{N}^2,\ k + l \le K_2\}$. Because the inter-arrival times, trip times and times to become a gift have exponential distribution, $(Y^N(t))$ is a Markov process, with finite state space

$$S^N = \Bigg\{y = (y_{1,j}, y_{2,k,l}, z)_{\{j\in\chi_1,(k,l)\in\chi_2\}}, y_{1,j} \in \frac{\mathbb{N}}{N_1}, y_{2,k,l} \in \frac{\mathbb{N}}{N_2}, z \in \frac{\mathbb{N}}{N}, \sum_{j\in\chi_1} y_{1,j} = 1,$$
$$\sum_{(k,l)\in\chi_2} y_{2,k,l} = 1, \sum_{j\in\chi_1} j\, y_{1,j} + \sum_{(k,l)\in\chi_2} (k+l)\, y_{2,k,l} + z \le M\Bigg\}.$$

The inequality in the previous definition of the state space S^N is due to the fact that the number of cars driving has be to added to the left-hand side of the inequality to obtain the total number M of cars in the system. Let us write its transitions from state $y \in S^N$. To simplify the notations, let us denote by

$$E_1 = \sum_{j\in\chi_1} j\, y_{1,j} \quad \text{and} \quad E_2 = \sum_{(k,l)\in\chi_2} (k+l)\, y_{2,k,l} \tag{1}$$

respectively the mean number of cars parked in cluster 1 and the mean number of cars plus gifts parked in cluster 2. Also, let us denote by $(e_{1,j}, e_{2,k,l}, e_0, j \in \chi_1,\ (k,l) \in \chi_2)$ the canonical basis of $\mathbb{R}^{|\chi_1|+|\chi_2|+1}$, where the cardinal of set A is denoted by $|A|$. The transitions, from state $y = (y_{1,j}, y_{2,k,l}, z) \in S^N$, are due to three events: a user arrival, a gift appearance or a car return. For example, when a user arrives at a station of cluster 2 with k cars and l gifts (for short of type $(2,k,l)$) to take a gift ($l>0$), the number of gifts decreases by 1. Since there are $y_{2,k,l}N_2$ possible stations, this happens at rate $\lambda_2 y_{2,k,l}N_2\mathbf{1}_{\{l>0\}}\big(p+(1-p)\mathbf{1}_{\{k=0\}}\big)$. Recall that p is the probability for a user arriving to a station in cluster 2 to choose a gift when cars and gifts are available. Thus the corresponding transition is the following.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k,l-1}-e_{2,k,l}) + \frac{e_0}{N} \quad \text{at rate} \quad \lambda_2 y_{2,k,l}N_2\mathbf{1}_{\{l>0\}}\big(p+(1-p)\mathbf{1}_{\{k=0\}}\big).$$

The other transitions are presented in the appendix. These transitions allow us to write the drift of process $(Y^N(t))_t$ which will be useful to the mean-field convergence (Proposition 2.1).

2. Mean-field limit

Our aim is to investigate the system when M, N_1 and N_2 get large at the same rate, for short, when N gets large. When N tends to $+\infty$, the process $(Y^N(t))$ given by the previous transitions converges in distribution to a deterministic function which is the unique solution of a given ODE. This result is given by the following proposition.

Proposition 2.1 (Mean-field convergence)**.** *For $T>0$, $(Y^N(t))_{t\in[0,T]}$ converges in distribution to $(y(t))_{t\in[0,T]}$ unique solution with $y(0)$ fixed of the following ODE*

$$\begin{aligned}
\frac{\mathbf{dy_{1,j}}}{\mathbf{dt}}(\mathbf{t}) &= y_{1,j+1}(t)\lambda_1\mathbf{1}_{\{j<K_1\}} + y_{1,j-1}(t)\left(\frac{c_1}{\alpha_1}\mu(s-\alpha_1E_1(t)-\alpha_2E_2(t)-z(t)) + \frac{q_1\mu_c}{\alpha_1}z(t)\right)\mathbf{1}_{\{j>0\}}\\
&- y_{1,j}(t)\left(\lambda_1\mathbf{1}_{\{j>0\}} + \frac{c_1}{\alpha_1}\mu(s-\alpha_1E_1(t)-\alpha_2E_2(t)-z(t))\mathbf{1}_{\{j<K_1\}} + \frac{q_1\mu_c}{\alpha_1}z(t)\mathbf{1}_{\{j<K_1\}}\right)\\
\frac{\mathbf{dy_{2,k,l}}}{\mathbf{dt}}(t) &= y_{2,k,l+1}(t)\lambda_2\mathbf{1}_{\{k+l<K_2\}}\big(p+(1-p)\mathbf{1}_{\{k=0\}}\big) + y_{2,k+1,l}(t)\lambda_2\mathbf{1}_{\{k+l<K_2\}}\big(1-p+p\mathbf{1}_{\{l=0\}}\big)\\
&+ y_{2,k+1,l-1}(t)\,\delta\,(k+1)\,\mathbf{1}_{\{k<K_2\}} + y_{2,k-1,l}(t)\left(\frac{c_2}{\alpha_2}\,\mu\,(s-\alpha_1E_1(t)-\alpha_2E_2(t)-z(t))\right.\\
&\left.+\frac{q_2\mu_c}{\alpha_2}z(t)\right)\mathbf{1}_{\{k>0\}} - y_{2,k,l}(t)\left(\lambda_2(1-\mathbf{1}_{\{k=0,l=0\}}) + \delta k + \frac{c_2}{\alpha_2}\mu(s-\alpha_1E_1(t)-\alpha_2E_2(t)\right.\\
&\left.- z(t))\mathbf{1}_{\{k+l<K_2\}} + \frac{(1-q)\mu_c}{\alpha_2}z(t)\mathbf{1}_{\{k+l<K_2\}}\right)\\
\frac{\mathbf{dz}}{\mathbf{dt}}(\mathbf{t}) &= -q_1\,\mu_c z(t)\sum_{j\in\chi_1} y_{1,j}(t)\,\mathbf{1}_{\{j<K_1\}} + \alpha_2\,\lambda_2\sum_{(k,l)\in\chi_2} y_{2,k,l}(t)\,\mathbf{1}_{\{l>0\}}\,\big(p+(1-p)\mathbf{1}_{k=0}\big)\\
&- q_2\mu_c z(t)\sum_{(k,l)\in\chi_2} y_{2,k,l}(t)\mathbf{1}_{\{k+l<K_2\}}.
\end{aligned} \tag{2}$$

Recall that, in these equations, s is the limiting number of cars per station and α_i the limiting proportion of stations in cluster i, $i\in\{1,2\}$.

The proof is standard (see [**?**]). The idea of the proof is that a Markov process can be written as the sum of a martingale term and a drift term in form of an integral on time. When N is large, one can prove that the process is tight. Moreover, the martingale term converges to 0. Then any limiting value satisfies an ODE. The uniqueness of the solution of the ODE gives the convergence of the process.

2.1. **The equilibrium point.** To investigate the steady-state behavior of the model, we study the equilibrium point $\bar{y}$ of the mean-field ODE written as follows

$$\frac{\mathbf{dy}}{\mathbf{dt}}(\mathbf{t}) = F(y(t))$$

where F comes from Proposition 2.1. It amounts to finding $\bar{y}$ such that

$$F(\bar{y}) = 0. \tag{3}$$

Note that the vector $\bar{y}$ is of dimension $1 + |\chi_1| + |\chi_2| = 1 + K_1 + K_2(1 + K_2)/2$. Finding a closed-form expression of the equilibrium point $\bar{y}$ is out of reach. Let us present two points of view: the first one is based on a nice queueing interpretation which holds for the no-gift case. The second is an analytic approach which should be relevant for the case with gifts but is beyond this work.

2.1.1. *The queueing interpretation for the no-gift case.* In this case, the existence and unicity of the equilibrium point $\bar{y}$ is proved. See [**?**] for details. In addition, $\bar{y}$ is given by a simple queueing interpretation of the mean-field limit. It gives that the limiting stationary number of cars at a station of cluster i, considered as a $M/M/1/K_i$ queue, has a geometric distribution $\nu_{\rho r_i, K_i}$ on $\{0, \dots, K_i\}$ with parameter ρr_i where for $i = 1, 2$, $r_i = \Lambda\mu\beta_i/\lambda_i$ with $\beta_i = q_i/\alpha_i$, $\Lambda = 1/\max_i(\mu\beta_i/\lambda_i)$ and ρ is the unique solution of the fixed point equation

$$s = \rho\Lambda + \sum_{i=1}^{2} \alpha_i m(\nu_{\rho r_i, K_i}). \tag{4}$$

In the previous equation, we denote by $m(\nu_{\rho,K})$ the mean of the geometric distribution $\nu_{\rho,K}$ on $\{0, \dots, K\}$ with parameter ρ, given by

$$m(\nu_{\rho,K}) = \begin{cases} \frac{K}{2} & \text{if } \rho = 1 \\ \frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}} & \text{otherwise} \end{cases} \tag{5}$$

because, for $\rho = 1$, $\nu_{\rho,K}$ is the uniform distribution on $\{0, \dots, K\}$. It shows that the multidimensional equilibrium point equation (3) amounts to fixed point equation (4) on $\mathbb{R}_+$. This is the purpose of [**?**, Theorem 1] for the cluster case detailed in [**?**, Section 2.3].

2.1.2. *Analytical method for the model with gifts.* Taking into account the gift policy induces a change of classes between normal cars and gifts. This complicates considerably the search for an equilibrium point and changes the nature of the limiting objects involved. The question of existence and uniqueness of a solution of the equilibrium point equation (3) remains open. For simplicity, let us take the case $p = q_1 = 1$ in order to highlight the main difficulties of this problem. Remembering that $p = 1$ means that, when available, a gift is always chosen over a car in a station of normal zone, and $q_1 = 1$ means that all gifts are returned at a station of cluster 1. Heuristically, looking for an equilibrium point $\bar{y}$ means that the right-hand term in the mean-field ODE (2) is null. But since with obvious notations $\bar{y} = (\bar{y_1}, \bar{y_2}, \bar{z})$, note first that the set of $z(t)$ moving gifts is considered as a $M/M/\infty$ queue whose fluid limit is

$$\bar{z} = \frac{\alpha_2\lambda_2}{\mu_c}\frac{1 - \bar{y}_{2,.,0}}{1 - \bar{y}_{1,S}} \tag{6}$$

where $\bar{y}_{1,S}$ is the probability that a station in cluster 1 is saturated and $\bar{y}_{2,.,0}$ the probability that a station in cluster 2 has no gift, i.e. $1 - \bar{y}_{1,S} = \sum_{j\in\chi_1} \bar{y}_{1,j}\mathbf{1}_{\{j<K_1\}}$ and $1 - \bar{y}_{2,.,0} = \sum_{(k,l)\in\chi_2} \bar{y}_{2,k,l}\mathbf{1}_{\{l>0\}}$. Then a queuing interpretation similar to that for the no-gift case holds.

Indeed, at equilibrium, a station of cluster 1 can be considered as a $M/M/1/K_1$ queue, with arrival rate

$$\bar{\gamma_1} = \frac{1}{\alpha_1}\left(c_1\mu(s-\alpha_1\bar{E}_1-\alpha_2\bar{E}_2-\bar{z})+q_1\mu_c\bar{z}\right) \tag{7}$$

where $\bar{E}_i$ are defined by (1) and service rate λ_1. It is well known that its invariant measure is a geometric distribution on $\{0,\dots,K_1\}$ with parameter $\bar{\rho}_1=\bar{\gamma}_1/\lambda_1$, i.e. $\bar{y}_{1,j}=\rho_1^j(1-\rho_1)/(1-\rho_1^{K+1})$ for $0\le j\le K_1$. Note that, plugging equation (6) into (7), $\bar{\rho}_1$ depends on $\bar{y}$, only by $\bar{y}_1$ and $\bar{y}_2$. Moreover $\bar{y}_2=\pi_{\bar{\rho}_2,K_2}$ where

$$\bar{\rho}_2 = \frac{1}{\lambda_2\alpha_2}\left(c_2\mu(s-\alpha_1\bar{E}_1-\alpha_2\bar{E}_2-\bar{z})+q_2\mu_c\bar{z}\right)$$

and, for fixed ρ, π_{ρ,K_2} is the invariant measure of the Markov process on χ_2 with matrix jump $Q_{\rho,K}$ given by its non-null non-diagonal terms

$$\begin{cases} Q_{\rho,K}(n,n-e_1) & =\lambda_2\mathbf{1}_{\{n_1>0\}}\\ Q_{\rho,K}(n,n+e_2) & =\lambda_2\rho\mathbf{1}_{\{n_2<K_2\}}\\ Q_{\rho,K}(n,n+e_1-e_2) & =\delta\mathbf{1}_{\{n_2>0\}}. \end{cases} \tag{8}$$

In conclusion the equilibrium point $\bar{y}$, solution of a multidimensional fixed point equation, can be expressed as a function of $(\bar{\rho}_1,\bar{\rho}_2)$ solution of a fixed point equation. It is summarized by the following result.

Proposition 2.2 (Equilibrium point). *An equilibrium point of the ODE is given as*

$$\bar{y} = \left(\nu_{\bar{\rho}_1,K_1},\pi_{\bar{\rho}_2,K_2},\frac{\alpha_2\lambda_2}{\mu_c}\frac{1-\sum_{k=0}^{K_2}\pi_{\bar{\rho}_2,K_2}(k,0)}{\sum_{k=0}^{K_1-1}\nu_{\bar{\rho}_1,K_1}(k)}\right)$$

where $\nu_{\bar{\rho}_1,K_1}$ is the geometric distribution on $\{0,\dots,K_1\}$ with parameter $\bar{\rho}_1$, $\pi_{\bar{\rho}_2,K_2}$ the invariant measure associated to $Q_{\bar{\rho}_2,K_2}$ given by (8) and $(\bar{\rho}_1,\bar{\rho}_2)$ is solution of the fixed point equation

$$\rho_i = \frac{1}{\lambda_i\alpha_i}\left(c_i\mu(s-E)+(q_i\mu_c-c_i\mu)\frac{\alpha_2\lambda_2}{\mu_c}\frac{1-\sum_{k=0}^{K_2}\pi_{\rho_2,K_2}(k,0)}{\sum_{k=0}^{K_1-1}\nu_{\rho_1,K_1}(k)}\right),\quad i\in\{1,2\} \tag{9}$$

with $E=\alpha_1E_1+\alpha_2E_2$, E_1 and E_2 being the means associated to ν_{ρ_1,K_1} and π_{ρ_2,K_2}.

Proposition 2.2 gives that the problem of existence and uniqueness of the equilibrium point $\bar{y}$ amounts to that of $(\bar{\rho}_1,\bar{\rho}_2)$ i.e. of the solution of fixed point equation (9). It is beyond the scope of this study. Nevertheless, in this direction, one can wonder if a closed-form expression can be found for $\pi=\pi_{\rho,K}$. Let $\gamma_2=\lambda_2\rho$. The global balance equation associated to π is

$$\begin{aligned}\pi_{k,l}&(\gamma_2+\lambda_2(1-\mathbf{1}_{\{k=l=0\}})+\delta k)\\ &=\mathbf{1}_{\{k+l<K_2\}}(\pi_{k,l+1}\lambda_2+\pi_{k+1,l}\lambda_2\mathbf{1}_{\{l=0\}})+\pi_{k+1,l-1}\delta(k+1)\mathbf{1}_{\{l>0\}}+\pi_{k-1,l}\gamma_2\mathbf{1}_{\{k>0\}}.\end{aligned} \tag{10}$$

An analytical approach could be based (see [**?**] for details) on generating function

$$F(x,y)=\sum_{(k,l)\in\chi_2}\pi_{(k,l)}x^ky^l.$$

The global balance equation (10) yields to a functional equation. Although the capacity K_2 is assumed to be finite throughout the hole paper, we present here this function equation for the case $K_2 = +\infty$ by sake of simplicity

$$F(x,y)\left(\gamma_2(1-x)+\lambda_2\left(1-\frac{1}{y}\right)\right) = F'_x(x,y)\delta(y-x)+\pi_{0,0}\lambda_2\left(1-\frac{1}{x}\right)+f(x)\lambda_2\left(\frac{1}{x}-\frac{1}{y}\right)$$

where $f(x)=\sum_{k=0}^{K_1}\pi_{k,0}x^k$.

3. Performance

In order to evaluate the impact of the incitative algorithm on the system behavior, a usual performance metric is used, i.e. the proportion of stations with no vehicle (car or gift) or no parking space available, called *problematic stations*. It gives how far the system is unbalanced.

Definition 3.1 (Performance Metric)**.** *Let $\bar{y}$ be the equilibrium point of the mean-field ODE obtained by Proposition 2.1. The performance metric is the limiting stationary proportion Pb of problematic stations given by*

$$Pb = \alpha_1(\bar{y}_{1,0}+\bar{y}_{1,K_1})+\alpha_2\left(\bar{y}_{2,0,0}+\sum_{k=0}^{K_2}\bar{y}_{2,k,K_2-k}\right)$$

where K_i is the station capacity and α_i the limiting proportion of stations for cluster i, $i\in\{1,2\}$.

The first sum into brackets is the proportion of empty and saturated stations in clusters 1, the first term $\bar{y}_{1,0}$ of stations with no car, the second term $\bar{y}_{1,K_1}$ of saturated stations in high-demand zone. The second sum into brackets is the proportion of empty and saturated stations in cluster 2, $\bar{y}_{2,0,0}$ of stations with neither cars nor gifts and $\sum_{k=0}^{K_2}\bar{y}_{2,k,K_2-k}$ of saturated stations in normal zone.

Optimizing the proportion of problematic stations means maximizing the number of transactions and the number of satisfied users. Our aim is to compare the performance with gifts and without gifts. The idea is to vary the fleet size parameter s, which is the limiting ratio of the total number of cars M by the total number of stations N, in order to analyze how much flexibility the gift policy gives to an operator who wants to increase the fleet size without harming the system.

3.1. Analysis of the model without gifts.

From Section 2.1.1, the proportion of problematic stations Pb in this case is given by

$$Pb=\sum_{i=1}^{2}\alpha_i\frac{1-\rho r_i}{1-(\rho r_i)^{K_i+1}}(1+(\rho r_i)^{K_i+1})$$

where $\alpha_i=\lim_N N_i/N$. For $i=1,2$, the proportion of problematic stations in cluster i as a function of s is given by the parametric curve

$$\rho\mapsto\left(\rho\Lambda+\sum_{i=1}^{2}\alpha_i m(\nu_{\rho r_i,K_i}),\frac{1-\rho r_i}{1-(\rho r_i)^{K_i+1}}(1+(\rho r_i)^{K_i+1})\right)$$

where the first term $(1-\rho r_i)/(1-(\rho r_i)^{K_i+1})$ is the proportion of empty stations in cluster i and the second term $(\rho r_i)^{K_i+1}(1-\rho r_i)/(1-(\rho r_i)^{K_i+1})$ is the proportion of saturated stations in cluster i. As explained in Section 5.2 of [**?**], the proportion of problematic stations in cluster

i has a minimum $2/(K_i+1)$ for ρr_i equal to 1 i.e. for $\rho = 1/r_i$. Thus, plugging in equation (4), this minimum corresponds to

$$s_i^* = \frac{\Lambda}{r_i} + \sum_{i'=1}^{2} \alpha_{i'} m(\nu_{r_{i'}/r_i, K_{i'}}).$$

where $m(\nu_{\rho,K})$ is defined by equation (5). With the notations of the paper, it gives the following result.

Proposition 3.1 (Optimal performance per cluster without gift policy)**.** *For the model without gifts, the limiting stationary proportion of problematic stations in cluster $i \in \{1,2\}$ is minimal and equal to $2/(K_i+1)$ when*

$$s = s_i^* = \alpha_i \left(\frac{K_i}{2} + \frac{\lambda_i}{\mu q_i} \right) + \alpha_{3-i} \left(\frac{\gamma_{3-i}}{1-\gamma_{3-i}} - \frac{(K_{3-i}+1)\gamma_{3-i}^{K_{3-i}+1}}{1-\gamma_{3-i}^{K_{3-i}+1}} \right)$$

where $\gamma_{3-i} = (q_i\lambda_i\alpha_i)/(q_{3-i}\lambda_{3-i}\alpha_{3-i})$. The last term in brackets must be replaced by $K_{3-i}/2$ for $\gamma_{3-i} = 1$.

Note that, for $s = s_i^*$ which minimizes the proportion of problematic stations in cluster i, the proportion of problematic stations in cluster $i' \neq i$ is not optimal and is exactly $\nu_{r_{i'}/r_i, K_{i'}}(0) + \nu_{r_{i'}/r_i, K_{i'}}(K_{i'})$. Thus minimizing the problematic stations in both clusters simultaneously is not possible.

For values of Figure 2b and $\alpha_1 = \alpha_2 = 0.5$, Proposition 3.1 gives $s_1^* = 29.9$ and $s_2^* = 13.1$, and for $\alpha_1 = 0.28$ and $\alpha_2 = 0.72$, $s_1^* = 21.9$ and $s_2^* = 20.4$, which can be checked in Figure 2.

Note the U-shape of the curves plotted in Figure 2b. This shape is typical of these performance curves (cf [**?**]). Indeed, for small values of the mean number of cars per pseudo-station, the proportion of empty stations is large and close to 1. Similarly, if the mean number of cars per station is large, the proportion of saturated stations is large and close to 1. Since the performance criterion includes both cases, the U-shape is observed. The contribution of empty and saturated stations to the proportion of problematic stations is illustrated by [**?**, Figure 2] where the proportions of empty, saturated and problematic stations are plotted.

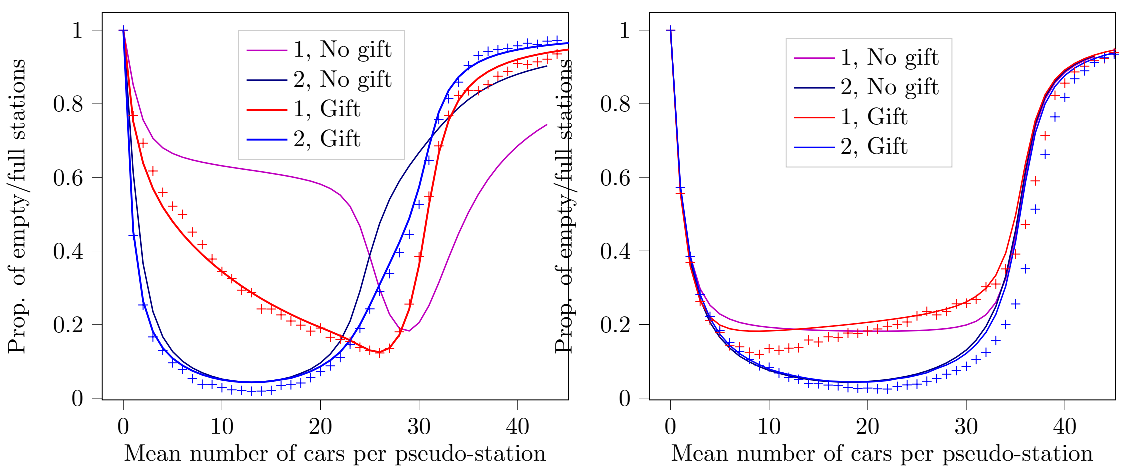
3.2. **Numerical solution.** First of all, we obtain numerically the equilibrium point $\bar{y}$ of the mean-field ODE established in Proposition 2.1, solution of the fixed point equation (3), as a function of the fleet size parameter s. There are many tools to solve such equation. We use the Anderson method implemented in Scipy, a Python library.

Figure 2a plots the performance Pb numerically obtained as a function of the fleet size parameter s, for the two-cluster model with and without gifts for a naive case: both clusters have the same number of stations, so that the ratios $\alpha_1 = \alpha_2 = 0.5$, and everyone follows the gift policy. That means the probability p that a user picks up a gift if gifts and cars are available and the probability q_1 a gift is returned to cluster 1 are equal to 1. All other parameters are given in Figure 2. Figure 2a shows that, for cluster 2, the cases with and without gifts are similar. But, for cluster 1, for this set of parameter values, it seems that an efficient gift policy ($p = q_1 = 1$) would allow an operator to increase the fleet size without harming the system performance and even with improving it. Indeed, for a whole range of values of the fleet size parameter s, typically $s \leq 20$, the high demand zone suffers from a lack of available cars. About 60% of the stations in the high demand zone are empty for a fleet size parameter s between 10 and 20. The effect of the incentive policy is significant in this case, since the proportion of empty stations in cluster 1 falls under 40% and even reaches 20% for $s = 20$.

Note that the crosses are simulations of the system with $N_1 = 50$ and $N_2 = 50$, the other parameters are given in Figure 2. Compared to the performance curves obtained numerically, it validates that the mean-field limit provides a good approximation for N_1 and N_2 large enough.

Figure 2b plots the performance numerically obtained for the two-cluster model with and without gifts for a more realistic case. The number of stations in the high demand zone is significantly smaller than in the normal one, the ratios are respectively $\alpha_1 = 0.28$ and $\alpha_2 = 0.72$. Figure 2b shows that the performance curves fit for small and large parameter fleet size s for both cases, with and without gifts. In between, there is a plateau where the proportion of problematic stations is close to its minimum. This implies that varying fleet size parameter s around its optimum does not degrade too much the performance which remains close to its optimum. This stability is important for the operator. The minimum proportion of problematic stations should depend on capacities K_1 and K_2, user arrival and trip rates. It is remarkable that the two plateaux correspond to the same values of s. Thus, the stations in cluster 1 do not saturate for s smaller than 30. Despite their small capacity, the high demand in cluster 1 limits the saturation.

In addition, Figure 2b shows that, for a small s, the gift policy slightly improves the performance. It is true until the two curves intersect at $s \simeq 12$. Above this value, on the plateau of cluster 1, the performance is slightly worse with the gift policy. Indeed, gifts seem to saturate cluster 1 and this slightly decrease the system performance. The mean-field approximation is again validated by simulation. See the crosses curve.



(A) $N_1 = N_2 = 50$

(B) $N_1 = 28$, $N_2 = 72$

Figure 2. Performance for both clusters (1 for the high-demand zone, and 2 for the normal zone) is numerically computed from equilibrium point equation as a function of the fleet size per station in a system with and without gifts, compared with the simulation curve in crosses. $K_1 = 15$, $K_2 = 45$, $\lambda_1 = 2.6$, $\lambda_2 = 1$, $\mu = \mu_c = 0.65$, $\delta = 1/14$, $c_1 = 0.5$ and $p = q_1 = 1$.

4. Discussion

4.1. **Discussion of the model.** *Discrete Markovian framework.* The exponential distributions are assumed to obtain a Markov discrete state process, i.e. the number of gifts and cars in the different stations. It is not true in real systems. It seems true for inter-arrival times of users at a station, but not trip times which seems heavy-tailed due to some very long trips. The behavior of the system can be affected by a log-normal trip time distribution compared to an exponential one. The threshold is deterministic. Intuitively, we assume that exponential distribution with the same mean for the threshold does not change the behavior of the network. Large stochastic networks with general service time distributions are still largely unexplored. We stay in a convenient framework.

Space-homogeneity. In order to simplify the presentation, we assume that parameters do not depend on the stations. This mean-field approach can be extended to a completely heterogeneous model. It is out of the scope of the paper.

Time-homogeneity. In real systems, some parameters, like the arrival rate of users, depend on time. The mathematical model does not take this into account.

Reservation. In real car-sharing systems, cars can be booked. It seems that such a study can still be conducted.

4.2. **Future work.** The analysis highlights an interesting random walk in the quarter-plane. Its study seems necessary to obtain further analytical results. Another model seems also necessary to obtain analytically the proportion of gifts in the system, to see the price to pay by the operator to implement such a policy.

The authors would like to thank Communauto for funding and allowing to do this study. They also thank the Natural Science and Engineering Research Council of Canada (NSERC) for funding.

Appendix A. Transitions of the empirical measure process

In this section, we present the detailed transitions of the empirical measure process $(Y^N(t))$ introduced in Section 1.4. The transitions, from state $y = (y_{1,j}, y_{2,k,l}, z) \in S^N$, are given by

- **User arrival.**
  - A user arrival at a station in cluster 2 with k cars and l gifts (for short of type $(2, k, l)$) taking a gift.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k,l-1} - e_{2,k,l}) + \frac{e_0}{N} \quad \text{at rate} \quad \lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{l>0\}}\big(p + (1-p)\mathbf{1}_{\{k=0\}}\big).$$

  - A user arrival at a station of type $(2, k, l)$ taking a normal car.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k-1,l} - e_{2,k,l}) \quad \text{at rate} \quad \lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{k>0\}}\big(1 - p + p\mathbf{1}_{\{l=0\}}\big).$$

  - A user arrival at a station of type $(1, j)$.

$$y \longrightarrow y + \frac{1}{N_1}(e_{1,j-1} - e_{1,j}) \quad \text{at rate} \quad \lambda_1 y_{1,j} N_1 \mathbf{1}_{\{j>0\}}.$$

- **Gift appearance.**
  - A car becoming a gift at a station of type $(2, k, l)$.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k-1,l+1} - e_{2,k,l}) \quad \text{at rate} \quad \delta k N_2 y_{2,k,l}.$$

- **Car return.**
  - A normal car returned at a station of type $(1, j)$.

$$y \longrightarrow y + \frac{1}{N_1}(e_{1,j+1} - e_{1,j}) \quad \text{at rate} \quad c_1 y_{1,j} \mu\Big(M - (E_1 N_1 + E_2 N_2 + zN)\Big)\mathbf{1}_{\{j<K_1\}}.$$

– A normal car returned at a station of type $(2, k, l)$.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k+1,l} - e_{2,k,l}) \quad \text{at rate} \quad c_2 y_{2,k,l} \mu \Big(M - (E_1 N_1 + E_2 N_2 + zN) \Big) \mathbf{1}_{\{k+l<K_2\}}.$$

– A gift returned at a station of type $(1, j)$.

$$y \longrightarrow y + \frac{1}{N_1}(e_{1,j+1} - e_{1,j}) - \frac{e_0}{N} \quad \text{at rate} \quad q_1 y_{1,j} \mu_c z N \mathbf{1}_{\{j<K_1\}}.$$

– A gift returned at a station of type $(2, k, l)$.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k+1,l} - e_{2,k,l}) - \frac{e_0}{N} \qquad \text{at rate} \qquad q_2 y_{2,k,l} \mu_c z N \mathbf{1}_{\{k+l<K_2\}}.$$

INRIA Paris, DI-ENS, PSL Research University, Paris, France
E-mail address: christine.fricker@inria.fr, bianca.marin-moreno@polytechnique.edu

MODAL'X, UMR CNRS 9023, UPL, Université Paris Nanterre
E-mail address: hmrad@parisnanterre.fr

Mathematical and Industrial Engineering, Polytechnique Montreal, Montreal, Canada, H3T 1J4
E-mail address: amaury.philippe@polymtl.ca, martin.trepanier@polymtl.ca