Spatial load balancing for EV charging customers

Andres Ferragut

Universidad ORT Uruguay

Al4OPT Seminar - Georgia Tech - February 2023

Collaborators



Emiliano Espíndola



Diego Marvid



Prof. Fernando Paganini

Introduction

Electrical vehicle (EV) adoption is currently growing exponentially.

Less carbon emissions, noise and other efficiency benefits.

Problems:

- Charging is power and energy intensive for the network.
- Still, charging requires a lot of time.
- We need to build a charging infrastructure to replace gas stations.

• How to estimate charging demand in a region.

• How to design the public charging infrastructure.

• How to design incentives for users to reduce network congestion.

• What can we do to exploit flexibility in user demand to reduce power draw.

Spatial load balancing

Dynamic queueing model

Stability results

Simulations

Estimating spatially distributed demand

Final remarks

Spatial load balancing

Dynamic queueing model

Stability results

Simulations

Estimating spatially distributed demand

Final remarks



Spatially distributed requests





Minimum distance assignment, does it make sense?



Maybe it's better to balance more evenly ...

Define:

- \bar{q}_i cars arrive at location *i*.
- There are \overline{s}_j chargers at location *j*.
- *C* is the *cost matrix*, $c_{ij} \ge 0$ is the cost of getting from $i \rightarrow j$.

Define:

- \bar{q}_i cars arrive at location *i*.
- There are \overline{s}_j chargers at location *j*.
- *C* is the *cost matrix*, $c_{ij} \ge 0$ is the cost of getting from $i \rightarrow j$.

Then one would like to solve:

$$\min\sum_{i,j}c_{ij}\gamma_{ij},$$

subject to:

$$egin{aligned} &\sum_{j}\gamma_{ij}=ar{q}_{i}, \quad orall i, \ &\sum_{i}\gamma_{ij}=ar{s}_{j}, \quad orall j, \ &\gamma_{ij}\in\{0,1\}. \end{aligned}$$

Monge transport problem, Kantorovich relaxation...

It is easier to relax the integrality constraint:

$$\min_{\Pi \geqslant 0} \sum_{i,j} c_{ij} \pi_{ij},$$

subject to:

$$\sum_{j} \pi_{ij} = ar{q}_i, \quad orall i, \ \sum_{i} \pi_{ij} = ar{s}_j, \quad orall j.$$

Monge transport problem, Kantorovich relaxation...

It is easier to relax the integrality constraint:

$$\min_{\Pi \geqslant 0} \sum_{i,j} c_{ij} \pi_{ij},$$

subject to:

$$\sum_{j} \pi_{ij} = ar{q}_i, \quad orall i,
onumber \ \sum_{i} \pi_{ij} = ar{s}_j, \quad orall j.$$

Relaxation is exact if quantities are integers... [Paganini et al CDC 2022]

A simple example...

Assume $\bar{q}_i \equiv 1$ and $\bar{s}_j \equiv 1$, and the number of cars and chargers is the same, then:

$$\min_{\Pi \geqslant 0} \sum_{i,j} c_{ij} \pi_{ij},$$

subject to:

$$\sum_i \pi_{ij} = 1 \quad orall j, \quad \sum_j \pi_{ij} = 1, \quad orall i$$

i.e. Π is a doubly stochastic matrix.

The vertices of the feasibility region are exactly the permutation matrices (Birkhoff theorem), so the solution is integer valued.

Allocation problem

If instead supply is abundant $(\sum_j \bar{s}_j \ge \sum_i \bar{q}_i)$ then we can formulate:

Problem 1



subject to:

$$\sum_{j} \pi_{ij} = \bar{q}_i, \quad \forall i,$$

 $\sum_{i} \pi_{ij} \leqslant \bar{s}_j, \quad \forall j.$

Dual formulation

Let's write the Lagrangian of Problem 1:

$$L_1(\Pi,\mu) = \sum_{i,j} c_{ij}\pi_{ij} + \sum_j \mu_j \left[\sum_i \pi_{ij} - \bar{s}_j\right] = \sum_{i,j} (c_{ij} + \mu_j)\pi_{ij} - \sum_j \mu_j \bar{s}_j.$$

Now minimize over Π subject to $\sum_j \pi_{ij} = \bar{q}_i$:

$$\min_{\Pi}\sum_{i,j}(c_{ij}+\mu_j)\pi_{ij}-\sum_{j}\mu_j\bar{s}_j,$$

subject to:

$$\sum_j \pi_{ij} = ar q_i \quad orall i.$$

Remark: the problem decouples over *i*.

Solution structure of Problem 1

For each *i*, send the "mass" \bar{q}_i from location *i* to the "cheapest" station using the modified cost $c_{ij} + \mu_j$.

• μ_j naturally acts as a congestion price on station *j*.

Solution structure of Problem 1

For each *i*, send the "mass" \bar{q}_i from location *i* to the "cheapest" station using the modified cost $c_{ij} + \mu_j$.

• μ_j naturally acts as a congestion price on station *j*.

Example: Let c_{ij} be the Euclidean distance then:

- If supply \bar{s}_j is abundant, $\mu_j \equiv 0$ and the routing splits traffic along Voronoï cells.
- When congestion occurs, $\mu_j > 0$ and the attraction region of station *j* is such that:

$$d(x, y_j) + \mu_j \leqslant d(x, y_k) + \mu_k \quad \forall k \neq j$$

Examples

Gaussian distribution for demand, centered at the origin.





Dual function

Define now:

$$\gamma_i(C,\mu) = \arg\min_{z\in\Delta}\sum_{j=1}^n (c_{ij}+\mu_j)z_j.$$

where Δ is the unit simplex, then:

$$\pi_{ij}^*(C,\mu) = \bar{q}_i \gamma_{ij}(C,\mu),$$

and

$$D_1(\mu) = \sum_i \bar{q}_i \min_j (c_{ij} + \mu_j) - \sum_j \mu_j \bar{s}_j,$$

piecewise-linear and concave, with super-gradient

$$\partial_j D_1 = \sum_i \bar{q}_i \gamma_{ij} (c_{ij} + \mu_j) - \bar{s}_j.$$

Andres Ferragut, Univ. ORT Uruguay

Explicit congestion penalty

We may also use a soft constraint on \overline{s}_j :

Problem 2

$$\min_{\Pi \geqslant 0, s \geqslant 0} \sum_{i,j} c_{ij} \pi_{ij} + \sum_{j} \phi_j(s_j),$$

subject to:

$$\sum_{j} \pi_{ij} = ar{q}_i, \quad orall i,
onumber \ \sum_{i} \pi_{ij} = s_j, \quad orall j.$$

Here $\phi_j(\cdot)$, increasing and convex, $\phi_j(0) = 0$, measures congestion at node *j*.

Explicit congestion

Lagrangian duality

Problem 2 is always feasible, and extends the range to $s_j > \overline{s}_j$ which amounts to waiting for service. Its Lagrangian is:

$$L_{2}(\Pi, s, \mu) = \sum_{i,j} c_{ij} \pi_{ij} + \sum_{j} \phi_{j}(s_{j}) + \sum_{j} \mu_{j} \Big[\sum_{i} \pi_{ij} - s_{j} \Big]$$

= $\sum_{i,j} (c_{ij} + \mu_{j}) \pi_{ij} + \sum_{j} [\phi_{j}(s_{j}) - \mu_{j}s_{j}].$

Explicit congestion

Lagrangian duality

Problem 2 is always feasible, and extends the range to $s_j > \bar{s}_j$ which amounts to waiting for service. Its Lagrangian is:

$$egin{aligned} L_2(\Pi, s, \mu) &= \sum_{i,j} c_{ij} \pi_{ij} + \sum_j \phi_j(s_j) + \sum_j \mu_j \Big[\sum_i \pi_{ij} - s_j \Big] \ &= \sum_{i,j} (c_{ij} + \mu_j) \pi_{ij} + \sum_j \left[\phi_j(s_j) - \mu_j s_j
ight]. \end{aligned}$$

- The minimum over Π is the same as before.
- The minimum over *s* is the negative of the Fenchel transform of ϕ :

$$\phi_j^*(\mu_j) = \min_{s_j} \{\mu_j s_j - \phi_j(s_j)\}.$$

Dual function

The dual function becomes:

$$D_2(\mu) = \sum_i \bar{q}_i \min_j (c_{ij} + \mu_j) - \sum_j \phi_j^*(\mu_j).$$

If ϕ_j is increasing and differentiable, then the optimum should satisfy:

$$\phi_j'(s_j^*) = \mu_j,$$

and we have the following gradient:

$$\partial_j D_2(\mu) = \sum_i \bar{q}_i \gamma_{ij} (c_{ij} + \mu_j) - [\phi'_j]^{-1}(\mu_j).$$

Spatial load balancing

Dynamic queueing model

Stability results

Simulations

Estimating spatially distributed demand

Final remarks

The previous analysis considered a "one-shot" allocation, but in practice:

- Requests will arrive asynchronously.
- The system should generate incentive compatible congestion signals.
- If demand is stationary, the dynamics should converge to an equilibrium consistent with what a central planner would do.

Here is when queueing theory and fluid models come into play!

System model

- Each location *i* receives a flow of requests at rate $r_i(t)$ requests/sec.
- The load balancing rule (to be designed) splits input flows into rates $r_{ij}(t)$ from location *i* to station *j*.
- The total flow rate into station *j* is thus $a_j := \sum_i r_{ij}$.

System model

- Each location *i* receives a flow of requests at rate $r_i(t)$ requests/sec.
- The load balancing rule (to be designed) splits input flows into rates $r_{ij}(t)$ from location *i* to station *j*.
- The total flow rate into station *j* is thus $a_j := \sum_i r_{ij}$.
- The state variable *s_j*(*t*) reflects the current assignment of station *j*; if it exceeds capacity *s_j*, we allow the possibility of waiting for service.
- The state evolution corresponds to the fluid queue

$$\dot{s}_j = a_j - d_j(s_j),$$

where d_j is the departure rate from station j, $d_j(0) = 0$.

System model Adding congestion signals

Key idea

Use time as the common ground for transportation and congestion costs.

Key idea

Use time as the common ground for transportation and congestion costs.

- Let c_{ij} be the transport time from $i \rightarrow j$ (e.g. $\frac{1}{\nu}d(x_i, y_j)$)
- Let μ_j be the waiting time at server *j*.
- Then $c_{ij} + \mu_j$ is the total delay before service.
- Selfish drivers then split according to

$$r_{ij}(t) = r_i \gamma_{ij}(C, \mu(t))$$

with γ_{ij} as before.

A natural fluid model for the *waiting time* would be:

$$\mu_j := rac{[s_j - ar{ extsf{s}}_j]^+}{d_j(s_j)}$$

• $[s_j - \bar{s}_j]^+$ is the excess number of requests at station *j*, i.e. the queue.

Dividing by the departure rate we get the time to get service.

System dynamics

The complete dynamics in the occupation state varible s_i are then:

$$\dot{s}_j = \sum_i r_i \gamma_{ij}(C,\mu) - d_j(s_j),$$
 $\mu_j = rac{[s_j - ar{s}_j]^+}{d_j(s_j)}$

System dynamics

The complete dynamics in the occupation state varible s_i are then:

$$\dot{s}_j = \sum_i r_i \gamma_{ij}(C,\mu) - d_j(s_j),$$
 $\mu_j = rac{[s_j - ar{s}_j]^+}{d_j(s_j)}$

Remark: We still need to model the departure rate. We developed two different choices:

- Sojourn-time model \rightarrow customers have fixed sojourn times.
- Service-time model \rightarrow customers have fixed service requirements.

Spatial load balancing

Dynamic queueing model

Stability results

Simulations

Estimating spatially distributed demand

Final remarks

Sojourn time model

- Assume users have a *time budget* that must be split between reaching the station, waiting for service and getting service.
- In order to maximize service time, selfish users will seek to minimize $c_{ij} + \mu_j$.
- If *T* is the average sojourn time, the departure rate is:

$$d_j(t) = \frac{1}{T} s_j(t),$$

■ The queueing delay (congestion price) becomes:

$$\mu_j(s_j) = \frac{[s_j - \overline{s}_j]^+}{d_j(s_j)} = T \left[1 - \frac{\overline{s}_j}{s_j} \right]^+.$$

Sojourn time model

Congestion price:

$$\mu_j(s_j) = T \left[1 - \frac{\overline{s}_j}{\overline{s}_j} \right]^+.$$

Sojourn time model

Congestion price:



Penalty function:

Andres Ferragut, Univ. ORT Uruguay

Equilibrium

Let $\bar{q}_i = r_i T$ be the steady state number of customers in the system from location *i*, then:

Theorem

The following are equivalent:

(a) s* is an equilibrium point of the dynamics. In particular, with μ_j^{*} = φ'_j(s_j^{*}) there exists a split γ* ∈ γ(c, μ*) such that a_j^{*} = ∑_i r_iγ_{ij}^{*} = d_j(s_j^{*}). Set π_{ij}^{*} = q_iγ_{ij}^{*}.
(b) (π*, s*, μ*) is a saddle point of the Lagrangian L₂ of Problem 2.

Equilibrium

Let $\bar{q}_i = r_i T$ be the steady state number of customers in the system from location *i*, then:

Theorem

The following are equivalent:

(a) s* is an equilibrium point of the dynamics. In particular, with μ_j^{*} = φ'_j(s_j^{*}) there exists a split γ* ∈ γ(c, μ*) such that a_j^{*} = ∑_i r_iγ_{ij}^{*} = d_j(s_j^{*}). Set π_{ij}^{*} = q_iγ_{ij}^{*}.
(b) (π*, s*, μ*) is a saddle point of the Lagrangian L₂ of Problem 2.

So the distributed load balancing equilibrium solves:

$$\min_{\Pi \ge 0} \sum_{i,j} c_{ij} \pi_{ij} + \sum_{j} \phi_j(s_j),$$

with the preceding barrier function.

Stability

Theorem

The dual function $D_2(\mu)$ for Problem 2 is non-decreasing along trajectories of $\mu(t)$ arising from the sojourn-time dynamics.

Proof:

$$\begin{aligned} \frac{d}{dt}D_2(\mu(t)) &= \sum_j \frac{\partial D_2}{\partial \mu_j} \dot{\mu}_j = \sum_{j \in \mathcal{J}} \underbrace{\left[\sum_i r_i T \gamma_j (c_{ij} + \mu_j) - [\phi'_j]^{-1}(\mu_j)\right]}_{=Ta_j - s_j} \dot{\mu}_j \\ &= \sum_j \underbrace{\left[Ta_j - s_j\right]}_{=T\dot{s}_j} T \frac{\bar{s}_j}{s_j^2} \dot{s}_j = \sum_{j \in \mathcal{J}} \bar{s}_j \left(\frac{T\dot{s}_j}{s_j}\right)^2 \ge 0. \end{aligned}$$

Stability

Theorem

The dual function $D_2(\mu)$ for Problem 2 is non-decreasing along trajectories of $\mu(t)$ arising from the sojourn-time dynamics.

Proof:

$$\begin{split} \frac{d}{dt}D_2(\mu(t)) &= \sum_j \frac{\partial D_2}{\partial \mu_j} \dot{\mu}_j = \sum_{j \in \mathcal{J}} \underbrace{\left[\sum_i r_i T \gamma_j (c_{ij} + \mu_j) - [\phi'_j]^{-1}(\mu_j)\right]}_{=Ta_j - s_j} \dot{\mu}_j \\ &= \sum_j \underbrace{\left[Ta_j - s_j\right]}_{=T\dot{s}_j} T \frac{\bar{s}_j}{s_j^2} \dot{s}_j = \sum_{j \in \mathcal{J}} \bar{s}_j \left(\frac{T\dot{s}_j}{s_j}\right)^2 \ge 0. \end{split}$$

Combined with a LaSalle type of argument, we get asymptotic stability.

Service time model

- Assume users arrive at location *i* with rate *r*_{*i*}.
- They have a *service requirement*, with average service time *T*₀.
- The number of tasks in service at station *j* is $\min\{s_j, \bar{s}_j\}$ so the departure rate is:

$$d_j(t) = \frac{\min\{s_j, \bar{s}_j\}}{T_0}.$$

The congestion signal becomes:

$$\mu_j(\mathbf{s}_j) = \frac{[\mathbf{s}_j - \overline{\mathbf{s}}_j]^+}{d_j(\mathbf{s}_j)} = T_0 \left[\frac{\mathbf{s}_j}{\overline{\mathbf{s}}_j} - 1\right]^+.$$

Service time model

Complete dynamics

The complete dynamics are:

$$egin{aligned} \dot{s}_j &= \sum_i r_i \gamma_{ij} - d_j, \ d_j &= rac{1}{T_0} \min\{s_j, ar{s}_j\}, \ \mu_j &= T_0 \left[rac{s_j}{ar{s}_j} - 1
ight]^+, \ \gamma_{ij} &= rg\min_j \{c_{ij} + \mu_j\} \end{aligned}$$

Stability condition

Now the total service rate available is $\sum_{j} \bar{s}_{j}$.

So the natural stability condition is:

$$\sum_{i} \underbrace{r_i T_0}_{\bar{q}_i} \leqslant \sum_{j} \bar{s}_j$$

Stability condition

Now the total service rate available is $\sum_{j} \bar{s}_{j}$.

So the natural stability condition is:

$$\sum_{i} \underbrace{r_i T_0}_{\bar{q}_i} \leqslant \sum_{j} \bar{s}_j$$

This is exactly Problem 1 feasibility condition!

Service time model

Equilibrium

Theorem

Let $\bar{q}_i = r_i T_0$ and that the stability condition holds, then:

(a) If s^* is an equilibrium point of the dynamics, there exists a choice of split $\gamma^* \in \gamma(c, \mu^*)$ such that $a_j^* = \sum_i r_i \gamma_{ij}^* = d_j(s_j^*)$. Set $\pi_{ij}^* = \overline{q}_i \gamma_{ij}^*$; then (π^*, μ^*) is a saddle point of the Lagrangian L_1 .

(b) Let (π^*, μ^*) be a saddle point of the Lagrangian L_1 of Problem 1. Define:

$$s_{j}^{*} = \begin{cases} \sum_{i} \pi_{ij}^{*} & \text{if } \mu_{j}^{*} = 0; \\ \overline{s}_{j} \left(1 + rac{\mu_{j}}{T_{0}}
ight) & \text{if } \mu_{j}^{*} > 0; \end{cases}$$

Then s^{*} *is an equlibrium of the dynamics.*

Remark: Note the difference between the equilibrium occupation s^* per station, and the allocation of Problem 1, \hat{s} . The latter will always be bounded by capacity, whereas the former will exceed it at congested stations.

Remark: Part (b) requires condition $\sum_j \bar{s}_j \ge \sum_i \bar{q}_i$ for feasibility of Problem 1 and therefore existence of equilibria. Substituting $\bar{q}_i = r_i T_0$ we arrive at the stability condition stated prior to the theorem.

Service time model

Stability

The dynamic analysis in this case strongly parallels that of the previous problem.

Proposition

The dual function $D_1(\mu)$ for Problem 1 with $\bar{q}_i = r_i T_0$ is non-decreasing along trajectories.

Proof:

$$\begin{split} \frac{d}{dt}D_1(\mu(t)) &= \sum_j \frac{\partial D_2}{\partial \mu_j} \dot{\mu}_j = \sum_j \underbrace{\left[\sum_i r_i T \gamma_j (c_{ij} + \mu_j) - \bar{s}_j\right]}_{=T_0 a_j - s_j} \dot{\mu}_j \\ &= \sum_{j \in \mathcal{J}} [T_0 a_j - \bar{s}_j] \frac{T_0}{\bar{s}} [a_j - \bar{s}_j] = \sum_{j \in \mathcal{J}} \frac{T_0^2}{\bar{s}_j} (\dot{s}_j)^2 \ge 0; \end{split}$$

Spatial load balancing

Dynamic queueing model

Stability results

Simulations

Estimating spatially distributed demand

Final remarks

We now present some simulations for the sojourn time model:

- Cars arrive as a Poisson process of intensity r = 4 EV/min. in the square $[0, 1] \times [0, 1]$.
- Spatial distribution is a centered Gaussian.
- Exponential sojourn times with T = 60 min.
- Average number of customers: rT = 240.
- Car speed is v = 0.1, so the maximum travel time is ≈ 7 min.

Example 1

t

- Stations located at the center of the square and on the vertices.
- $\bar{s}_i \equiv 60$ amounts to 300 chargers.
- The central location becomes congested due to high demand in the ungion
- region.
- Problem 2 equilibrium:

$$s_1^* = 64.4, \quad s_j^* = 42.4, \, j = 2, \dots, 5,$$

and the congestion price $\mu_1 = 4.1$ minutes.





Example 1



Time evolution for the stochastic system. Solid lines are the ODE.

Andres Ferragut, Univ. ORT Uruguay

Example 2: randomly chosen locations

- Again 5 stations, but located at random points in the square.
- Rest of the parameters as before.
- Problem 2 equilibrium:

$$s_1^* = 48.9, s_2^* = 61.8, s_3^* = 61.6,$$

 $s_4^* = 49.5, s_5^* = 12.3.$

• Two stations become congested (the ones near the center of the region). $\mu_2^* \approx 1.7$ and $\mu_3^* \approx 1.5$ according to Problem 2.



Attraction regions

Example 2: randomly chosen locations



Time evolution for the stochastic system. Solid lines are the ODE.

Andres Ferragut, Univ. ORT Uruguay

Open questions

Space is discretized in our model, but it works great for distributed arrivals.

- Show convergence of the underlying stochastic model to the fluid limit.
 - Should be farly straightforward, but devil is in the details...
- Diffusion approximations around equilibrium?
 - Switching equations pose an issue (the $\min\{s_j, \bar{s}_j\}$ term)...
- Consider abandonments, incoming vehicles, etc. etc.

Spatial load balancing

Dynamic queueing model

Stability results

Simulations

Estimating spatially distributed demand

Final remarks

Ongoing work...

Motivation:

• We would like to know how much energy will EV cars need and where.

■ Idea: use current consumption measures at gas stations.

• Convert to kilometers and then to energy using efficiencies.

Spatial estimation

Main problem

• Let $\mathcal{X} \subset \mathbb{R}^2$ be a region of the plane.

Assume you have a *demand density* $g(x) : \mathcal{X} \to \mathbb{R}^+$ that you want to estimate.

Problem:

You can only have an estimate of:

$$y_i = \int_{V_i} g(x) dx$$

where V_i is a cell associated to measurement site s_i .

Example: s_i are the gas stations, you only have access to total demand, V_i is the "attraction region" of the gas station.

• We use Gaussian Radial Basis Functions to approximate g(x).

Using a sum of squares loss:

$$\mathcal{L} = \sum_{i} \left[\int_{V_i} g(x; heta) dx - y_i
ight]^2$$

we can perform gradient descent.

 Nice Montecarlo trick to estimate the integrals, relates to stochastic gradient descent.

Spatial estimation

Radial basis functions





Spatial estimation

Radial basis functions





More detailed model

A second model distinguishes among requests:

- Requests come from a spatial Poisson process of intensity $\lambda(x)$.
- Each request needs some exponential amount of service of parameter ν .
- You only have access to:

$$y_i = \sum_n \sigma_n \mathbf{1}_{\{n \in V_i\}}$$

where σ_n are the (iid) individual requests.

• For the second model, there is a *hidden variable* which is the number of requests.

• Using ideas from hidden Markov models, we can perform a maximum likelihood estimation of both $\lambda(x)$ and ν .

• For the $\lambda(x; \theta)$ we use RBFs again.

Spatial load balancing

Dynamic queueing model

Stability results

Simulations

Estimating spatially distributed demand

Final remarks

• EVs are popping up, we have to prepare the infrastructure.

- We showed how a simple routing rule can handle congestion.
- Lagrangian duality decentralizes the problem using natural congestion signals.
- The algorithm is incentive compatible, so users' choices are natural.
- We also discussed how to estimate demand in order to build the correct infrastructure.

Thank you!

Andres Ferragut ferragut@ort.edu.uy http://aferragu.github.io