No need to rush

Dealing with deadlines in EV charging

Andrés Ferragut

Universidad ORT Uruguay

20th INFORMS Applied Probability Society Conference - Brisbane - 2019

Collaborators

Martin Zeballos

Prof. Fernando Paganini

Motivation

- In the near future, electrical vehicles (EVs) will become an energy-intensive load to the grid.
- We need to provision infrastructure to provide charging capacity.
- The power and energy requirements will be significant, but users may be flexible in the charging time.
- Problem: how to make a smart use of available resources.

In this talk

Main contribution: Mean field analysis of EV scheduling with deadlines.

Highlights:

- We formulate a queueing model for an EV parking lot.
- Analyze the behavior of typical policies through fluid limits (mean field).
- Discuss overload scenarios and the impact on fairness.
- Discuss non-deadline aware policies

Outline

Queueing model of an EV recharge system

A comprehensive fluid model for scheduling policies

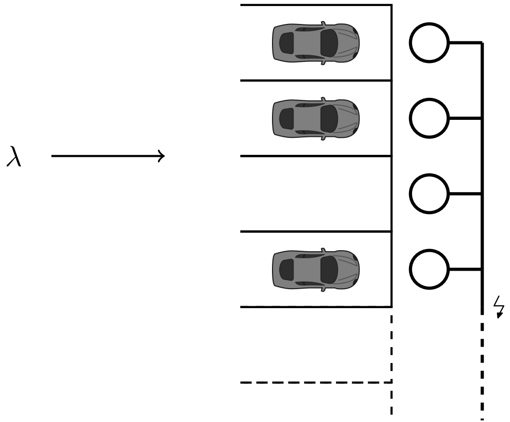
Deadline oblivious policies

Conclusions and future work

Outline

Parking lot model

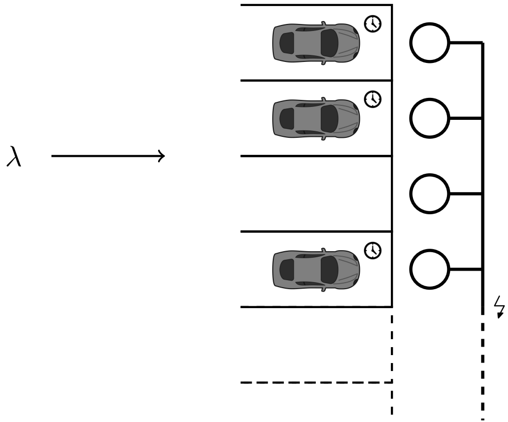
With individual charging stations



- λ = arrival rate of EVs.

Parking lot model

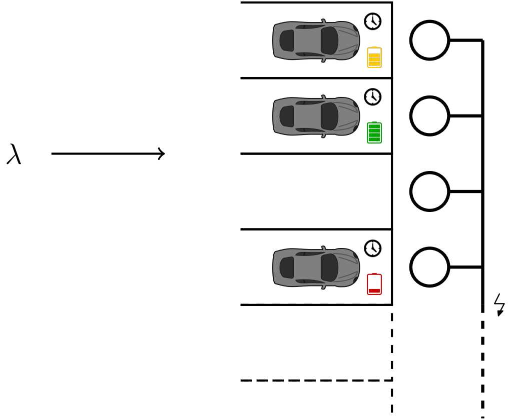
With individual charging stations



- λ = arrival rate of EVs.
- T_k = sojourn time (deadline).

Parking lot model

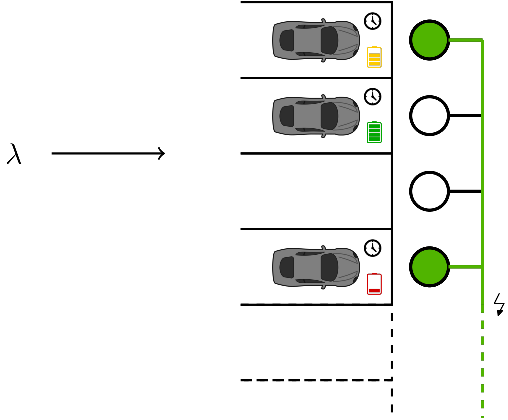
With individual charging stations



- λ = arrival rate of EVs.
- T_k = sojourn time (deadline).
- S_k = service time at nominal power.

Parking lot model

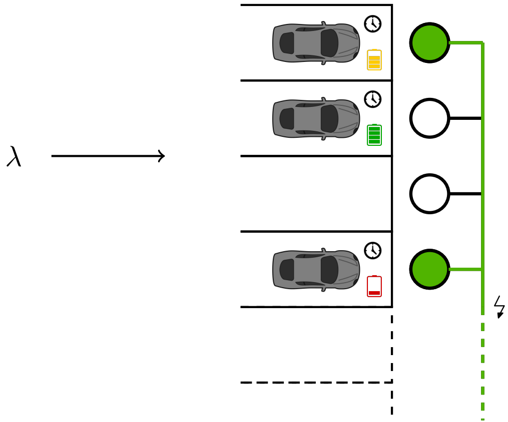
With individual charging stations



- λ = arrival rate of EVs.
- T_k = sojourn time (deadline).
- S_k = service time at nominal power.

Parking lot model

With individual charging stations

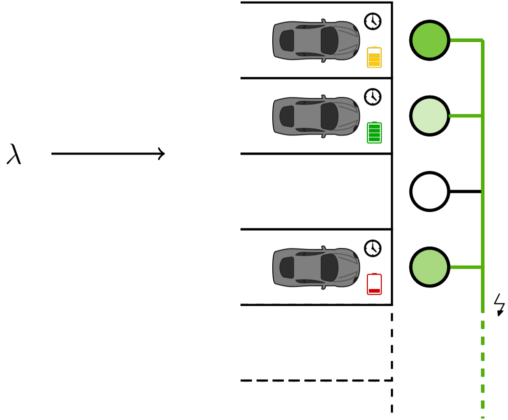


- λ = arrival rate of EVs.
- T_k = sojourn time (deadline).
- S_k = service time at nominal power.

- $\rho = \lambda \bar{S}$ is the load of the system.
- Interpretation: number of chargers that must be working, or average power requested.

Parking lot model

With individual charging stations



- λ = arrival rate of EVs.
- T_k = sojourn time (deadline).
- S_k = service time at nominal power.

- $\rho = \lambda\bar{S}$ is the load of the system.
- Interpretation: number of chargers that must be working, or average power requested.

- Remark: chargers may be individually controlled at intermediate levels.

Resource allocation problem

- Assume that parking space is unlimited...
- But we have a power budget C that we have to comply with.
- We may choose a charging rate $r_k(t)$ for each EV such that:

$$0 \leqslant r_k(t) \leqslant 1 \qquad \text{individual power constraint}$$

$$\sum_{k=1}^{N(t)} r_k(t) \leqslant C \qquad \text{system power constraint}$$

- We call a policy efficient if it does not loses charging opportunities (equality in at least one of the above).
- Moreover, we want each user to receive a fair share of service.

Outline

Queueing model of an EV recharge system

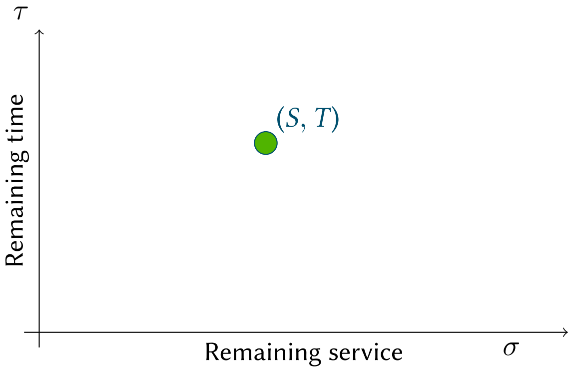
A comprehensive fluid model for scheduling policies

Deadline oblivious policies

Conclusions and future work

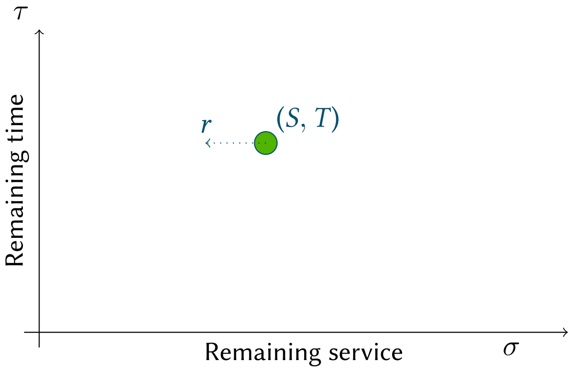
The service-sojourn time space

To visualize the problem, it is useful to consider the service-sojourn time space:



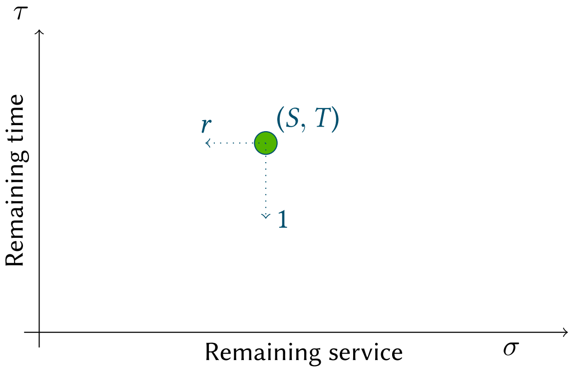
The service-sojourn time space

To visualize the problem, it is useful to consider the service-sojourn time space:



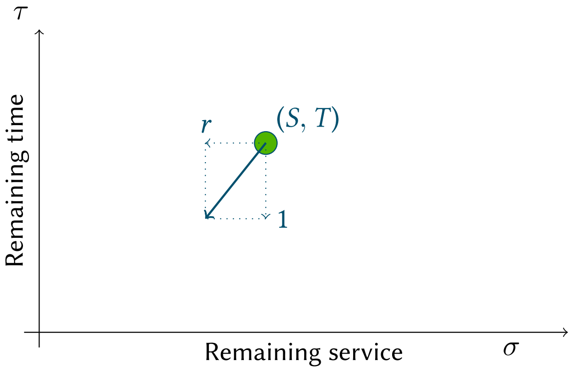
The service-sojourn time space

To visualize the problem, it is useful to consider the service-sojourn time space:



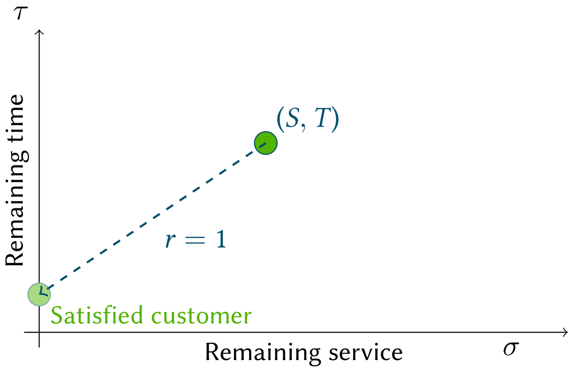
The service-sojourn time space

To visualize the problem, it is useful to consider the service-sojourn time space:



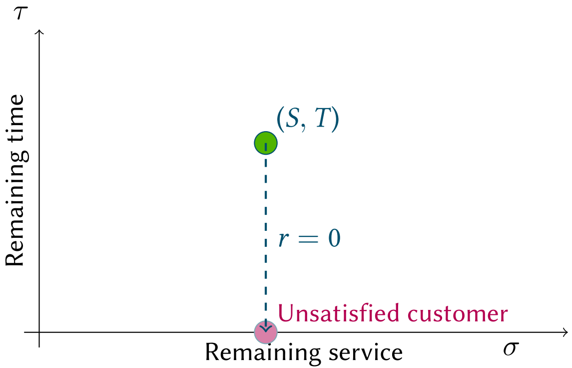
The service-sojourn time space

To visualize the problem, it is useful to consider the service-sojourn time space:



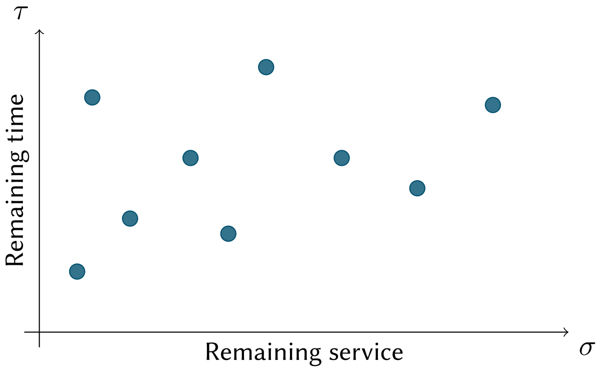
The service-sojourn time space

To visualize the problem, it is useful to consider the service-sojourn time space:



State descriptor

The state of the system is just the locations of points in this space:



State:

Point measure storing all points.,

$$\Phi_t = \sum_k \delta_{(\sigma_k(t)),\tau_k(t)}$$

State descriptor

The state of the system is just the locations of points in this space:

State:

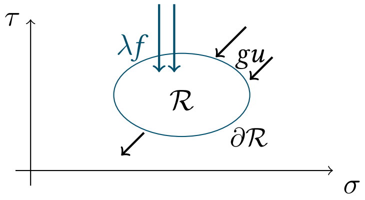
Point measure storing all points.,

$$\Phi_t = \sum_k \delta_{(\sigma_k(t)),\tau_k(t)}$$

Policy:

Vector field $\vec{u} = (-r, -1)$ that tells how each point moves.

If the system scale is large ($\lambda, C \to \infty$), we can treat the population as a fluid quantity $g(\sigma, \tau)$.



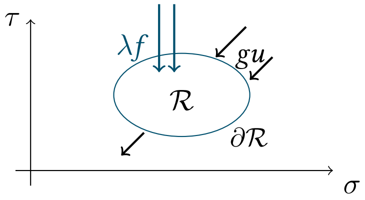
- λ = total arrival rate.
- $f(\sigma, \tau)$ joint density of (S, T).
- Charging policy:

$$\vec{u}(\sigma, \tau, g) = -\begin{bmatrix} r(\sigma, \tau, g) \\ 1 \end{bmatrix}.$$

Fluid model

A large scale limit of the system

If the system scale is large ($\lambda, C \to \infty$), we can treat the population as a fluid quantity $g(\sigma, \tau)$.



- λ = total arrival rate.
- $f(\sigma, \tau)$ joint density of (S, T).
- Charging policy:

$$\vec{u}(\sigma, \tau, g) = -\begin{bmatrix} r(\sigma, \tau, g) \\ 1 \end{bmatrix}.$$

Master fluid equation (using flow balance in a region):

$$\frac{\partial g}{\partial t} = \lambda f + \frac{\partial (rg)}{\partial \sigma} + \frac{\partial g}{\partial \tau}$$

Steady-state behavior of the fluid limit

- We are interested in the system equilibrium, i.e.:

$$\lambda f + \frac{\partial(rg)}{\partial\sigma} + \frac{\partial g}{\partial\tau} = 0.$$

- Fluid version of power constraints:

$$0 \leqslant r \leqslant 1; \qquad \iint rg\,d\sigma\,d\tau \leqslant C.$$

- Efficient policy: if charging opportunities are not wasted (equality in at least one of the above).

Remark: we can solve this explicitly for several policies!

The role of the load

Recall the load definition: $\rho = \lambda E[S]$.

Theorem

If $\rho < C$ (underload), then all efficient policies have the same fluid equilibrium with, $r \equiv 1$ (immediate service) and no reneged work.

The role of the load

Recall the load definition: $\rho = \lambda E[S]$.

Theorem

If $\rho < C$ (underload), then all efficient policies have the same fluid equilibrium with, $r \equiv 1$ (immediate service) and no reneged work.

Theorem

Si $\rho > C$ (overload), then all efficient policies have the same reneged work rate:

$$W = \int_0^\infty \sigma g(\sigma, 0) d\sigma = \rho - C.$$

The role of the load

Recall the load definition: $\rho = \lambda E[S]$.

Theorem

If $\rho < C$ (underload), then all efficient policies have the same fluid equilibrium with, $r \equiv 1$ (immediate service) and no reneged work.

Theorem

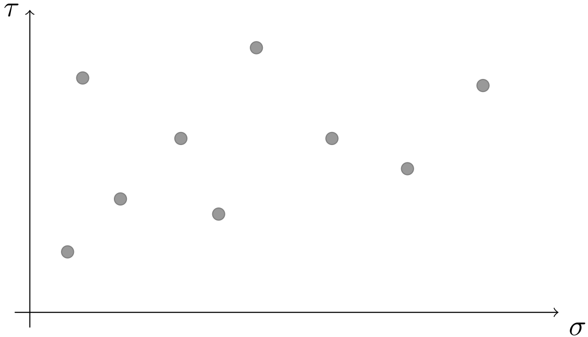
Si $\rho > C$ (overload), then all efficient policies have the same reneged work rate:

$$W = \int_0^\infty \sigma g(\sigma, 0) d\sigma = \rho - C.$$

Remark: The *distribution* of this reneged work is *highly dependent on the policy*, and may be unfair.

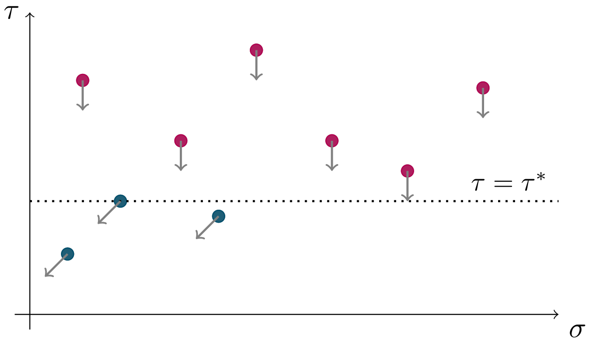
Example: Earliest Deadline First

- Idea: Prioritize EVs closer to departure.
- Long story in the processor scheduling community.
- Example: $N = 9, C = 3.$



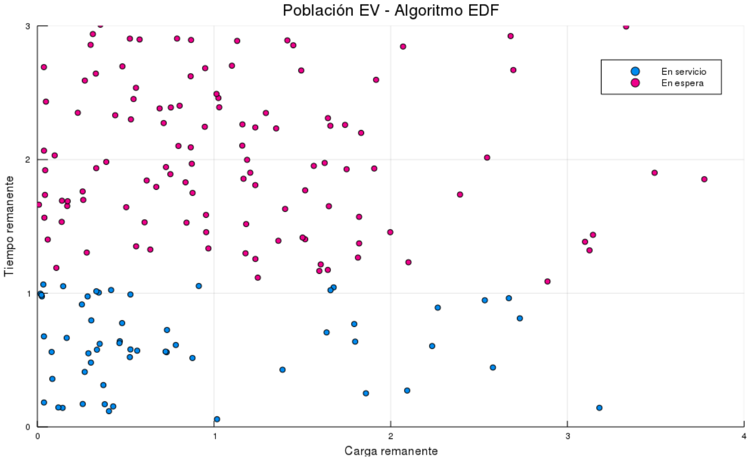
Example: Earliest Deadline First

- Idea: Prioritize EVs closer to departure.
- Long story in the processor scheduling community.
- Example: $N = 9$, $C = 3$.



Example: Earliest Deadline First

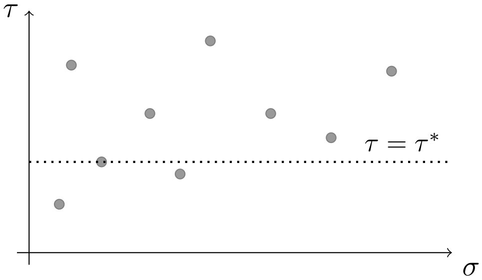
Large scale behavior



- $\rho = 100.$
- $C = 60.$
- $\lambda = 100$ ev/h.
- $\bar{S} = 1$h.
- $\bar{T} = 4$h.

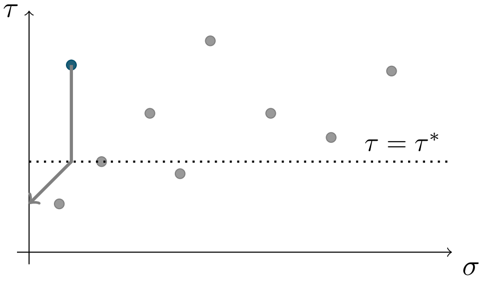
Earliest-Deadline-First (EDF)

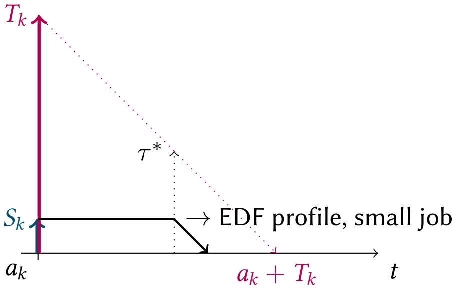
Large scale performance



Earliest-Deadline-First (EDF)

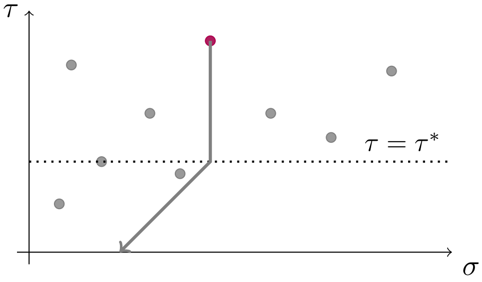
Large scale performance

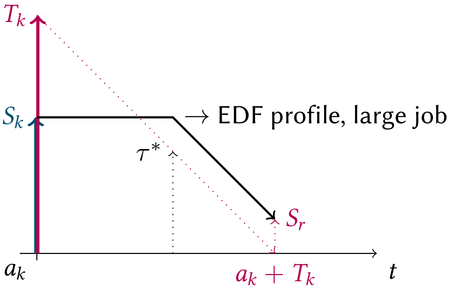




Earliest-Deadline-First (EDF)

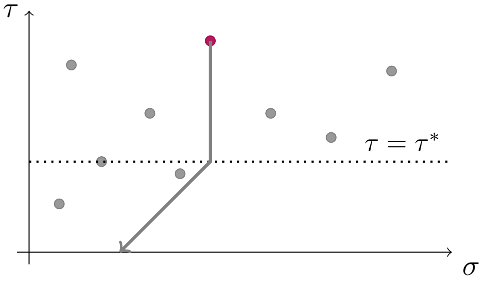
Large scale performance

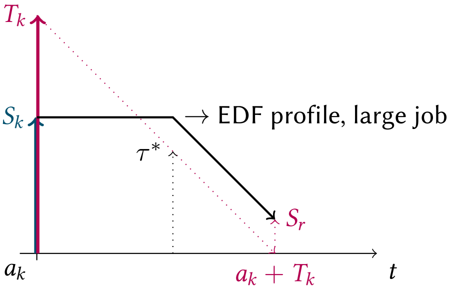




Earliest-Deadline-First (EDF)

Large scale performance





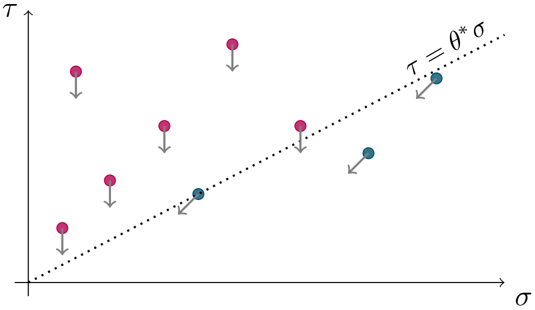
- The work received by each client is $S_a = \min\{S, \tau^*\}$.
- By imposing flow balance we can characterize τ^*:

$$\lambda E[\min\{S, \tau^*\}] = C$$

- It's unfair to large jobs.

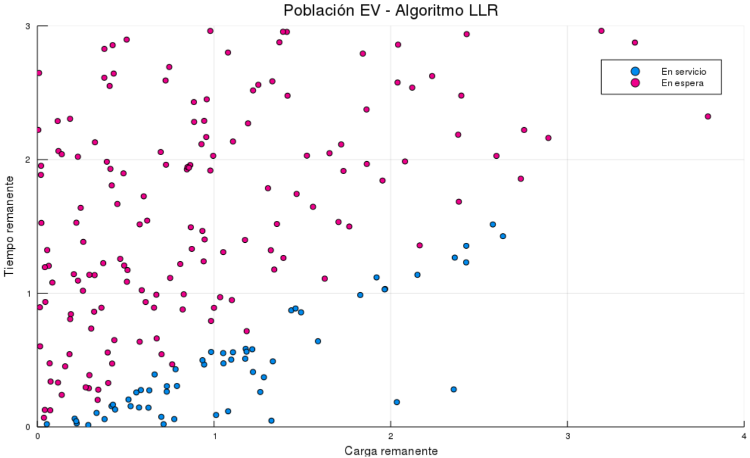
A fair policy: Least Laxity Ratio

- Idea: Prioritize EVs with greater relative urgency.
- Laxity ratio: $U_k = T_k / S_k$.
- Aims to balance between EDF and LLF.
- Example: $N = 9, C = 3$.



Least Laxity Ratio (LLR)

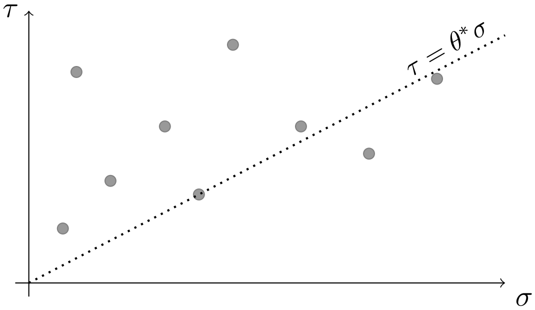
Example



- $\rho = 100.$
- $C = 60.$
- $\lambda = 100$ ev/h.
- $\bar{S} = 1$h.
- $\bar{T} = 4$h.

Least Laxity Ratio (LLR)

Performance



- In this case, the threshold is simply:

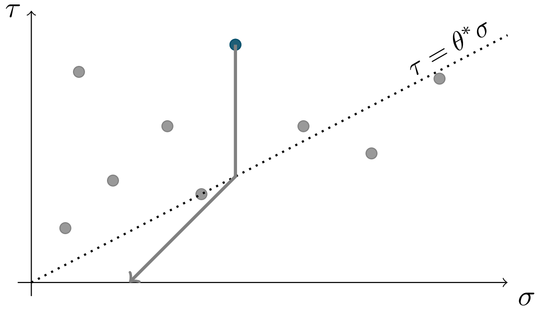
$$\lambda E[\theta^* S] = C \Rightarrow \theta^* = C/\rho.$$

- The work performed on each client is:

$$S_a = \theta^* S$$

Least Laxity Ratio (LLR)

Performance



- In this case, the threshold is simply:

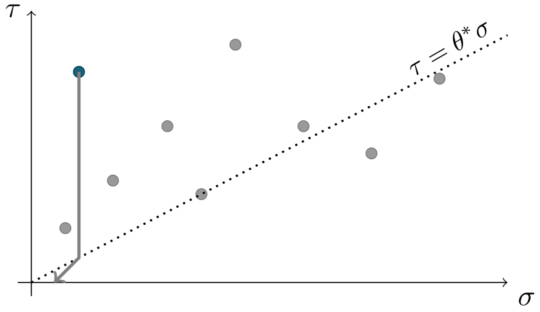
$$\lambda E[\theta^* S] = C \Rightarrow \theta^* = C/\rho.$$

- The work performed on each client is:

$$S_a = \theta^* S$$

Least Laxity Ratio (LLR)

Performance



- In this case, the threshold is simply:

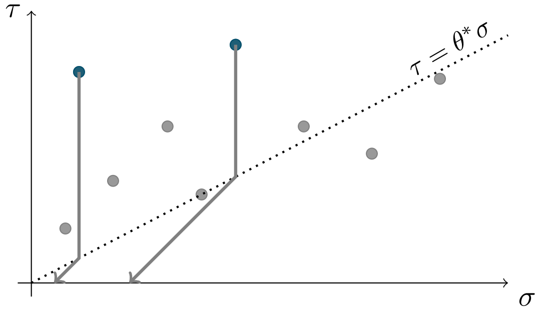
$$\lambda E[\theta^* S] = C \Rightarrow \theta^* = C/\rho.$$

- The work performed on each client is:

$$S_a = \theta^* S$$

Least Laxity Ratio (LLR)

Performance



- In this case, the threshold is simply:

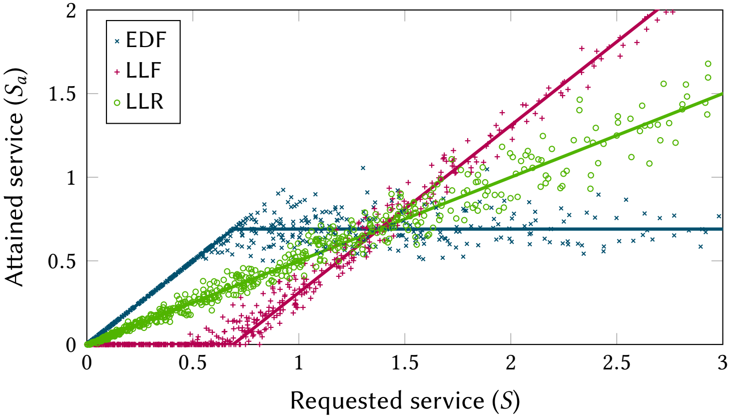
$$\lambda E[\theta^* S] = C \Rightarrow \theta^* = C/\rho.$$

- The work performed on each client is:

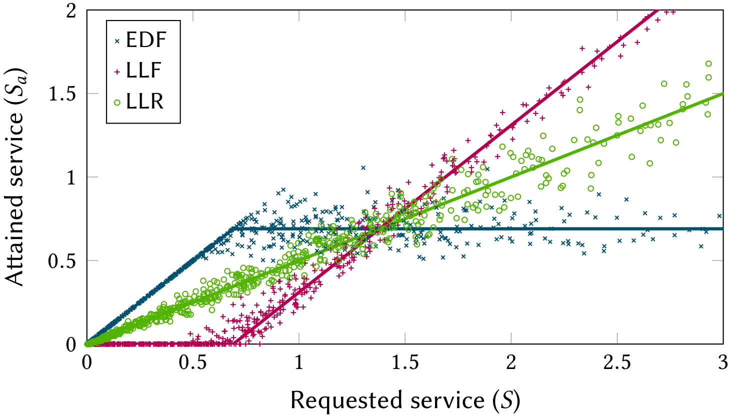
$$S_a = \theta^* S$$

- In overload, every vehicle gets the same relative service.

Comparison


EDF
LLF
LLR
Attained service (S_a)
Requested service (S)
0
0.5
1
1.5
2
2.5
3

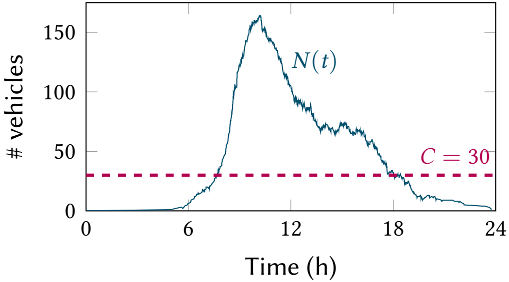
Comparison

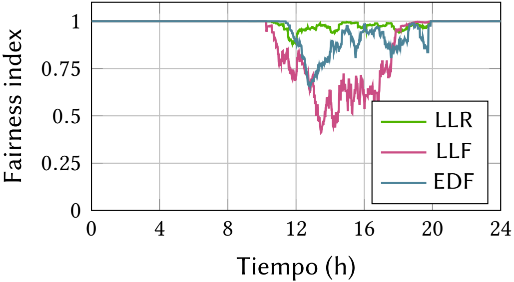


The LLR policy achieves a fair distribution of reneging

Application: a parking lot from California

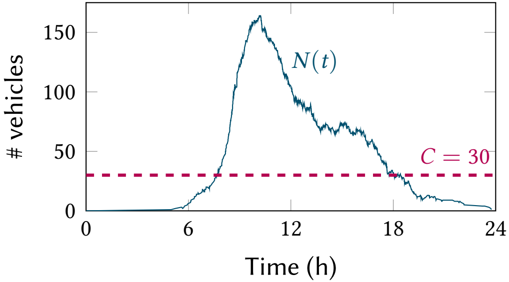
- We took data from a medium-sized employee parking lot in Silicon Valley.
- Tested the policies and computed Jain fairness index to measure fairness.

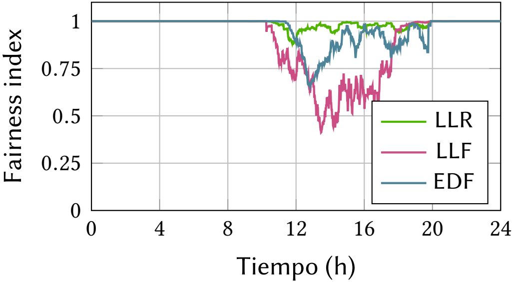




Application: a parking lot from California

- We took data from a medium-sized employee parking lot in Silicon Valley.
- Tested the policies and computed Jain fairness index to measure fairness.





- Our LLR algorithm preserves fairness even in time-varying conditions.
- The fluid model guide us in the algorithm design.

Outline

- Queueing model of an EV recharge system
- A comprehensive fluid model for scheduling policies
- Deadline oblivious policies
- Conclusions and future work

Deadline oblivious policies

- In practice, remaining service is easy to gauge (chargers are smart).
- However, deadline information is not easily available.

Deadline oblivious policies

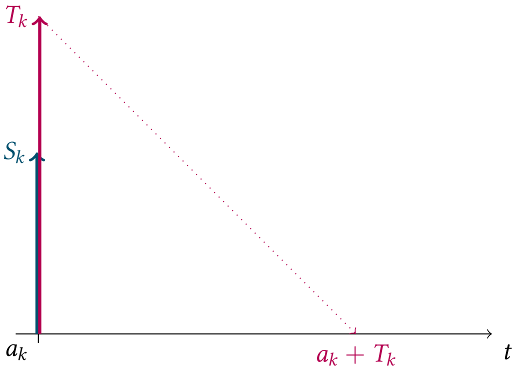
- In practice, remaining service is easy to gauge (chargers are smart).
- However, deadline information is not easily available.
- Question: can we emulate the behavior of deadline based policies without deadline information?

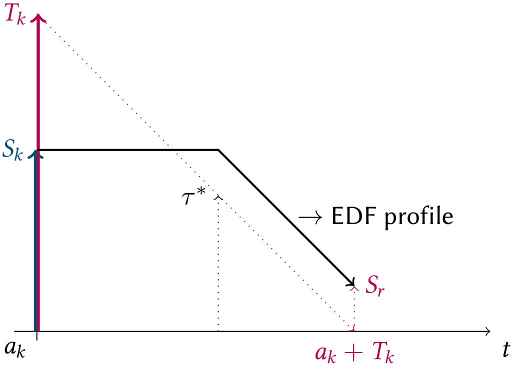
Deadline oblivious policies

- In practice, remaining service is easy to gauge (chargers are smart).
- However, deadline information is not easily available.
- Question: can we emulate the behavior of deadline based policies without deadline information?
- Idea: use the mean-field reasoning to gauge policy behavior.

Back to Earliest Deadline First

Mean field behavior





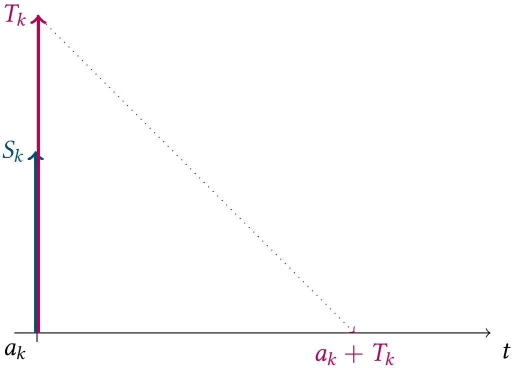
T_k
S_k
τ^*
$\rightarrow$ EDF profile
S_r
a_k
$a_k + T_k$
t

A really simple policy: Last Come First Served

- Really simple: just serve the users immediately upon arrival.
- Preemptive LCFS: new users interrupt the oldest ones.
- In practice: the more recent C jobs are served at any point in time.

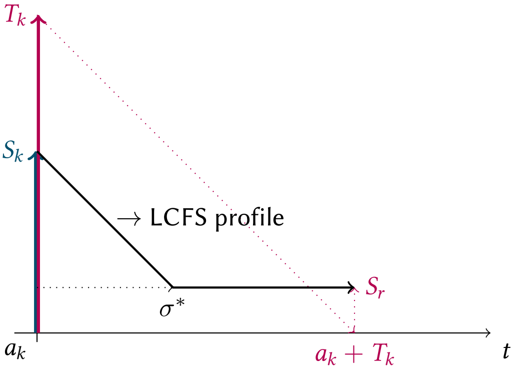
A really simple policy: Last Come First Served

Mean field behavior



A really simple policy: Last Come First Served

Mean field behavior



The threshold σ^* is the amount of time the load is served before being preempted for good.

LCFS vs EDF

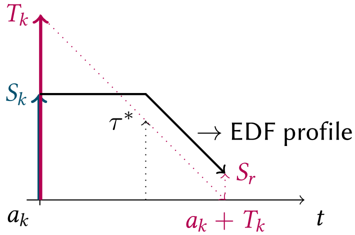
Proposition

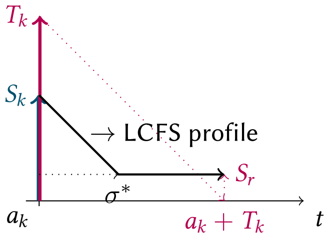
In the large scale limit, preemptive LCFS and EDF attain the same service.

LCFS vs EDF

Proposition

In the large scale limit, preemptive LCFS and EDF attain the same service.





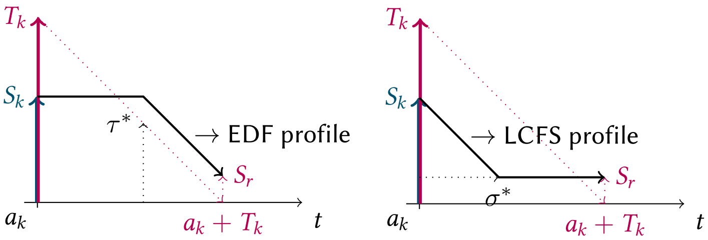
Applying flow balance, one should have $\sigma^* = \tau^*$ in overload.

$$S_a^{EDF} = S_a^{LLF} = \min\{S, \sigma^*\}$$

LCFS vs EDF

Proposition

In the large scale limit, preemptive LCFS and EDF attain the same service.



Applying flow balance, one should have $\sigma^* = \tau^*$ in overload.

$$S_a^{EDF} = S_a^{LLF} = \min\{S, \sigma^*\}$$

The Proposition extends to the Least-attained-service.

Least-Laxity-First alternative

With the same mean-field ideas we can find a proper substitute for Least-laxity-first:

Longest remaining processing time (LRPT)

- Serve the jobs with longest remaining service times.
- Preempt if a job arrives larger than the current ones.
- In practice: always serve the least served cars.

Least-Laxity-First alternative

With the same mean-field ideas we can find a proper substitute for Least-laxity-first:

Longest remaining processing time (LRPT)

- Serve the jobs with longest remaining service times.
- Preempt if a job arrives larger than the current ones.
- In practice: always serve the least served cars.

Proposition

In the large scale limit, LLF and LRPT attain the same service:

$$S_a = \max\{S - \sigma^*, 0\} \text{ where } \lambda E[(S - \sigma^*)^+] = C.$$

If $S < \sigma^*$, the job sees no service.

Even LLR can be substituted!

Consider the following policy

Least service ratio (LSR)

- Consider the jobs in increasing order of current σ/S.
- In practice: always serve the least served cars in proportion to their request.
- Non-local policy: depends on the original service time!

Even LLR can be substituted!

Consider the following policy

Least service ratio (LSR)

- Consider the jobs in increasing order of current σ/S.
- In practice: always serve the least served cars in proportion to their request.
- Non-local policy: depends on the original service time!

Proposition

In the large scale limit, LLR and LSR attain the same service:

$$S_a = \theta^* S \text{ where } \theta^* = C/\rho.$$

Outline

Queueing model of an EV recharge system

A comprehensive fluid model for scheduling policies

Deadline oblivious policies

Conclusions and future work

Conclusions

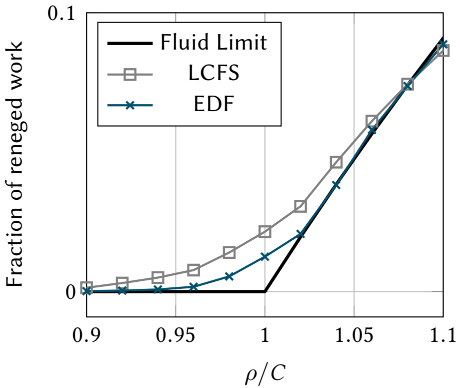
- We provided an unified framework to model EV charging policies. The mean field model yields insights on attained service for different policies.
- We used it to derive a new policy (least-laxity-ratio) that preserves some notion of fairness, even for time-varying scenarios.
- Using the mean-field paradigm, we analyzed non-deadline based policies, and showed strong connections between deadline-aware and deadline-oblivious equivalents.

Future work

Let me show you some graphs....

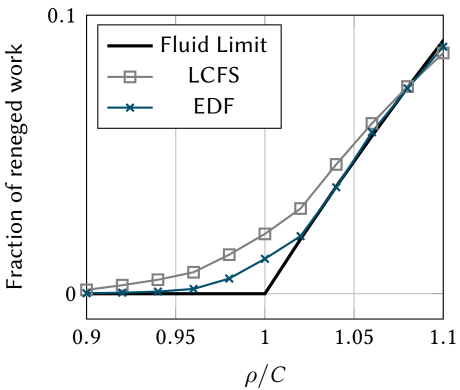
Future work

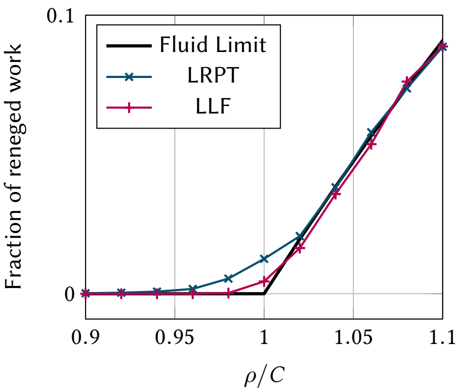
Let me show you some graphs....



Future work

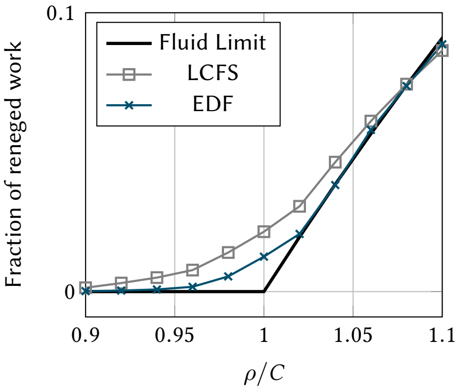
Let me show you some graphs....

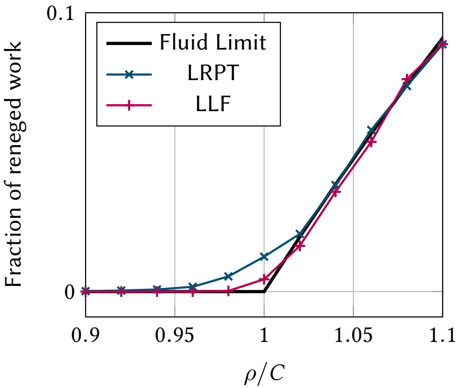




Future work

Let me show you some graphs....





Future work: The key is the transition region where diffusion approximations are needed.

Thank you!

Andrés Ferragut
ferragut@ort.edu.uy
http://fi.ort.edu.uy/mate
Paper: *Proportional fairness for EV charging in overload, IEEE Trans. on Smart Grid, 2019.*