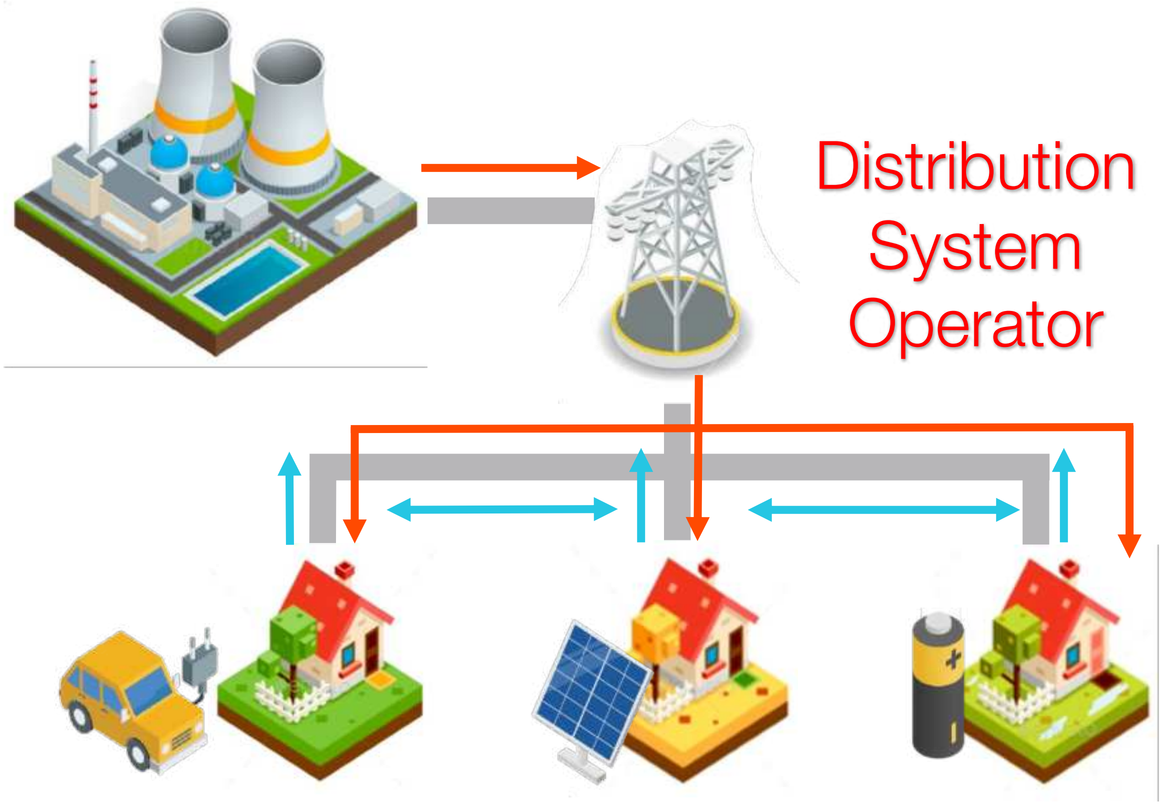
Sequential Decision Making for Optimal and Resilient Bi-Directional Integration of Electric Vehicles into Grid

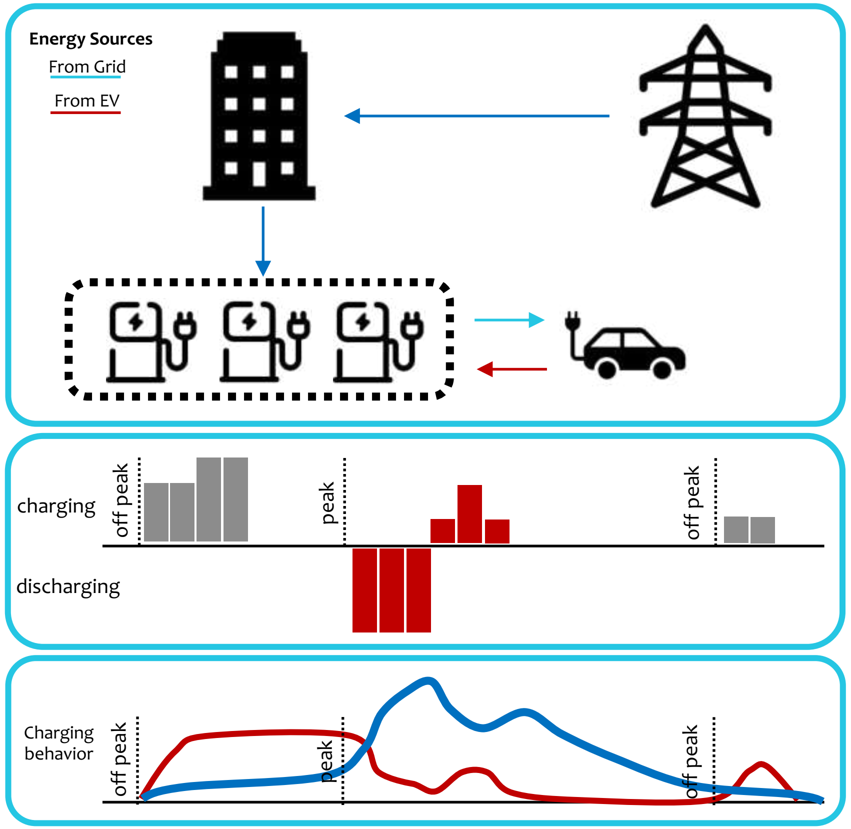
Abhishek Dubey, Vanderbilt University.

Funded in part by the National Science Foundation, Department of Energy and Nissan

Work done in Collaboration with Nissan North America

Distributed Energy Systems





- Move from **Centralized** generation and control to **distributed** generation and control enables **resilience**
- Allow consumers to be active participants and reduces load on current infrastructure.
- Integration of Electric Vehicles is changing the paradigm

VANDERBILT UNIVERSITY

Impact of Electrifying the Transportation Sector

Problem: **Integration of EV at scale including public transportation has led to newer challenges:** Managing the routes, charging infrastructure, and charging schedules of EV transit fleets such that operational costs are minimized, and the power grid is not overloaded by the additional power demand

Collaborative Optimization

Transportation Operators

Objective: minimize operational (charging) cost

Decision Variables:

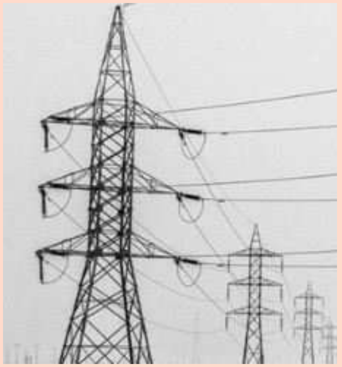
- Charging Locations
- Fleet composition (battery sizes, etc.)
- Route schedules
- Vehicle-to-route assignments
- Charging Schedule

Energy Utility

Objective: minimize grid impact

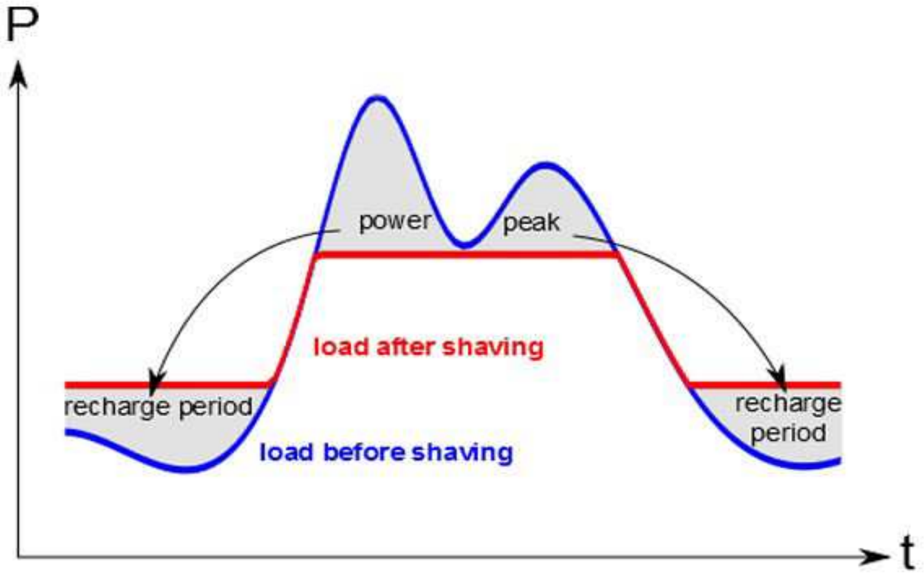
Decision Variables:

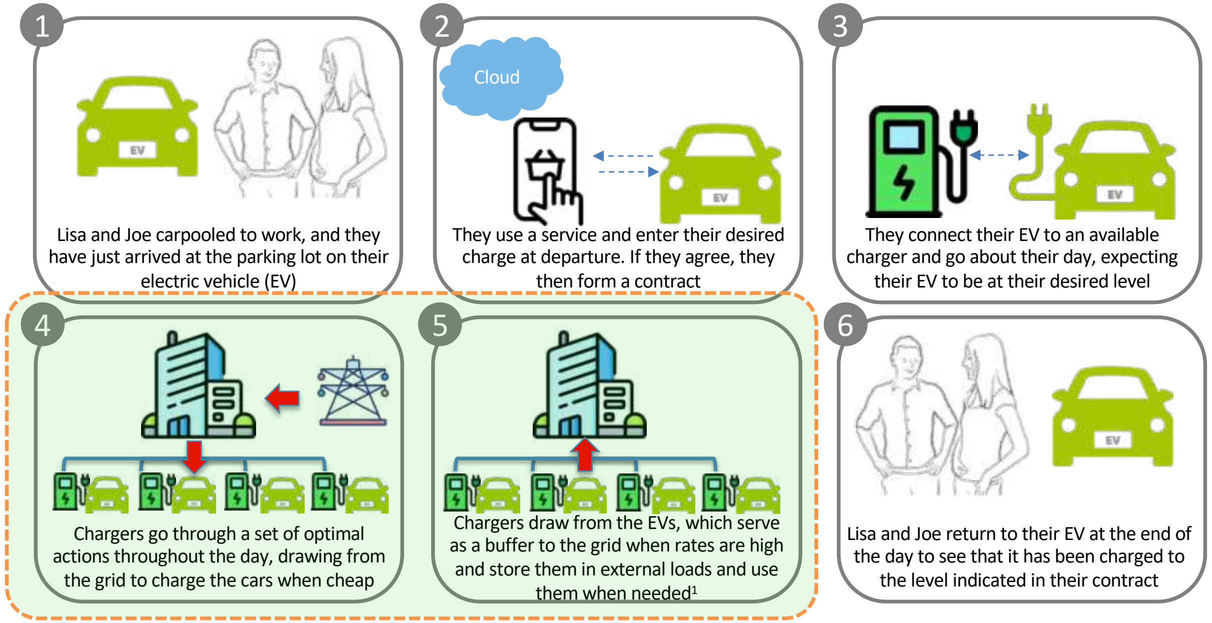
- Infrastructure upgrades
- Energy pricing structure



Challenges and Opportunities

- Increasing EV adoption is leading to **higher electricity demand** on building infrastructures.
- **Unmanaged EV charging** increases energy expenses and incurs peak demand charges.
- **Strategic V2B charging** optimizes energy usage, acting as **mobile batteries,** reducing costs and peak demand impacts.





[1]Supports energy sustainability and zero emissions buildings.

* from Peak shaving control method for energy storage by Karmiris *et. al*

Parts of the problem

- Cars:

- Arrival time
- Departure time
- Arrival SoC
- Required SoC
- Battery capacity

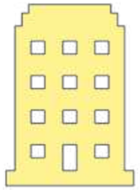
- Chargers:

- Directionality (Type)
- Min. charging/ discharging rate
- Max. charging/ discharging rate

- Power Utility company:

- Time-Of-Use rates
- Demand charge

- Building:



- Power usage profile
- Uncertain peaks

Solution Approach

Uncertainty Aware Systems

Mixed Linear Programming (MILP)

Pros:
- Optimal solution
- Fast convergence (for regular problem size)

Cons:
- Poor at handling uncertainty
- Needs complete information (Oracle)

MCTS: Monte-Carlo Tree Search

Pros:
- Fully online, anytime, and robust
- Robust to environmental changes

Cons:
- Slow, mainly due to the number of iterations
- Large action space is a major challenge

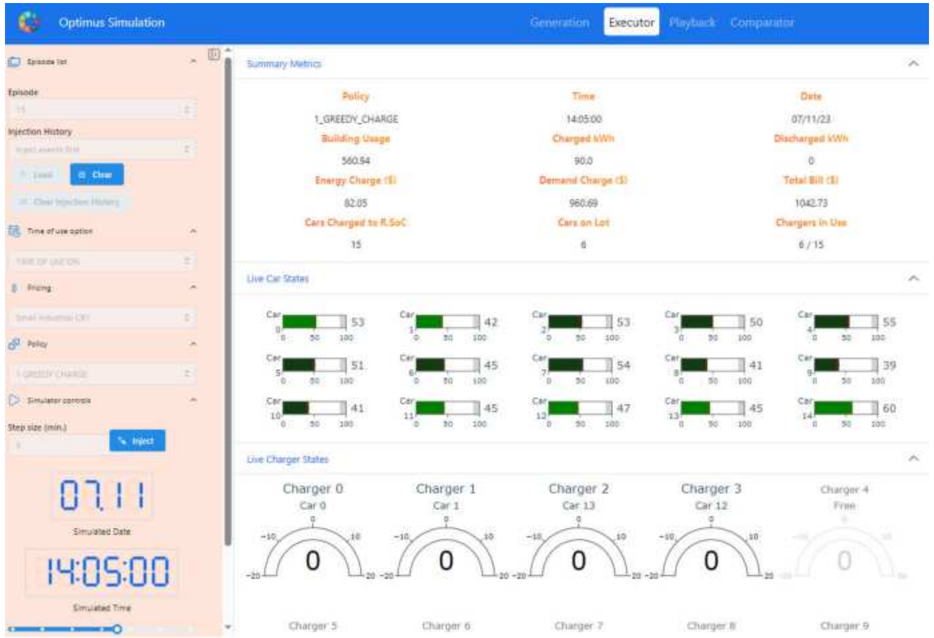
Reinforcement Learning System

Pros:
- Real-time decision-making.
- Adapts to uncertainties
- Effectively address complex input features
- Long-Term Optimization

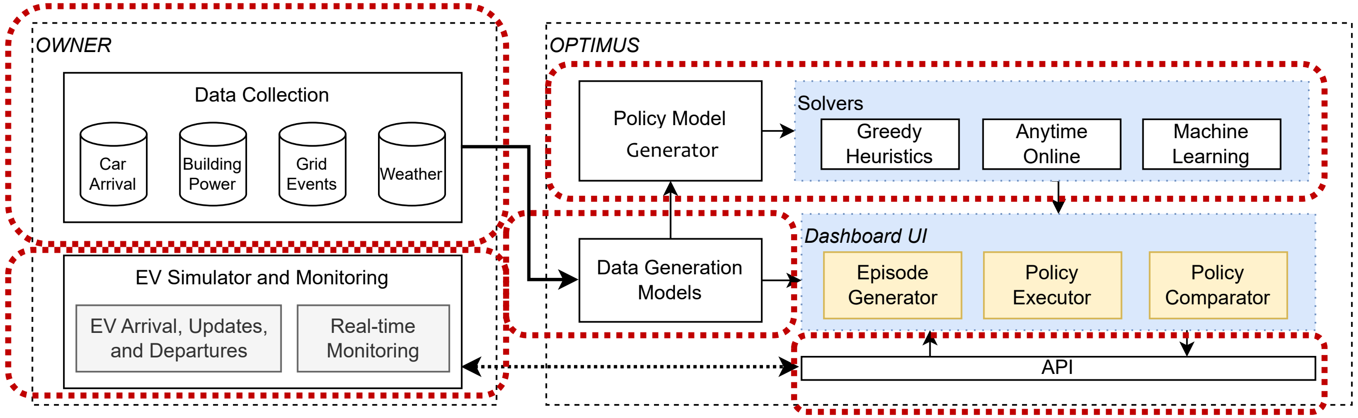
Cons:
- Challenges with large and continuous action spaces.
- Limited robustness to unseen conditions.

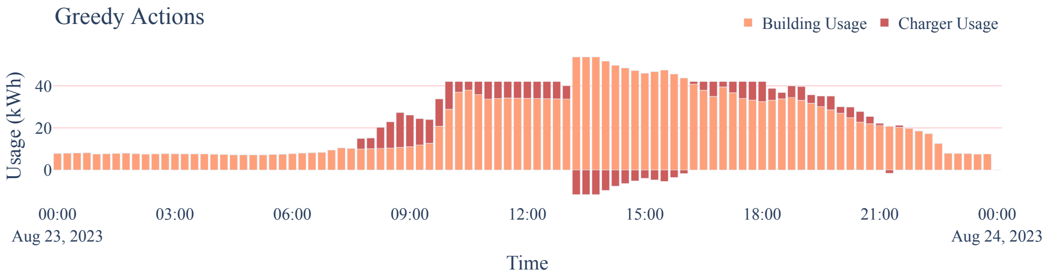
Policy Simulator

- End-to-end, discrete event-based simulator.
- Utilizes generative models to create realistic scenarios that match real-world use cases.
- Has several policies ranging from heuristics, optimal solvers, online solutions, and reinforcement learning-based approaches built-in.
- Features visualization with injection APIs for testing.



OPTIMUS
Supports REST APIs and acts as a digital twin



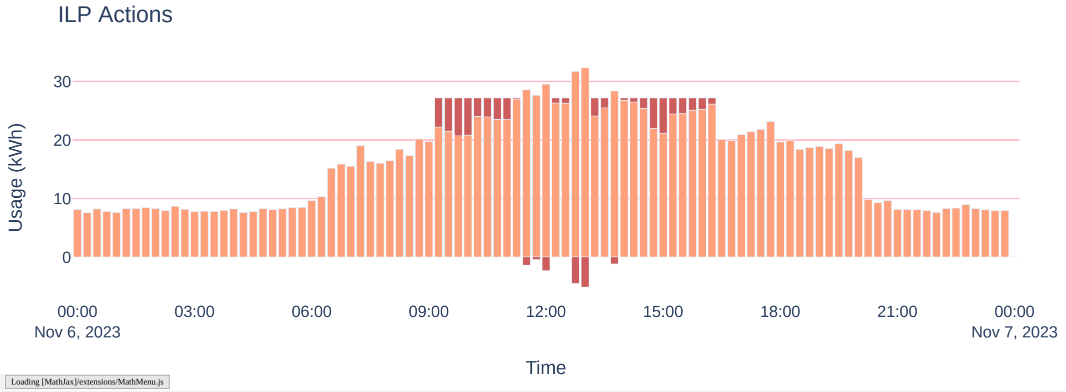


These depict the energy usage of both the chargers and the building. Actions for all chargers are aggregated at 15-minute intervals.

Optimal Solution: Mixed-Integer Linear Program (MILP)

■ When the future is known and events are deterministic, a MILP solver can give the optimal answer.

- We developed several variations of the solver, ranging from exact solutions to ones that are more adaptable to uncertainty and changes in the environment.
- Objective is to minimize the total bill.
- Constraints include:
  - Every car is assigned to a charger.
  - Cars remain connected until departure.



Opt-V2B:

- Uses **exact arrival and departure times** of cars.
- Required SoC at departure is a constraint

Opt-V2B-TW:

- **Allows flexible arrivals, departures** - assumes a time window of arrival and departure for each car (like, +/- 15 minutes) around the exact times

Opt-V2B-pen:

- **Allows any time departures** - Deviation from required SoC is a penalty, not a constraint

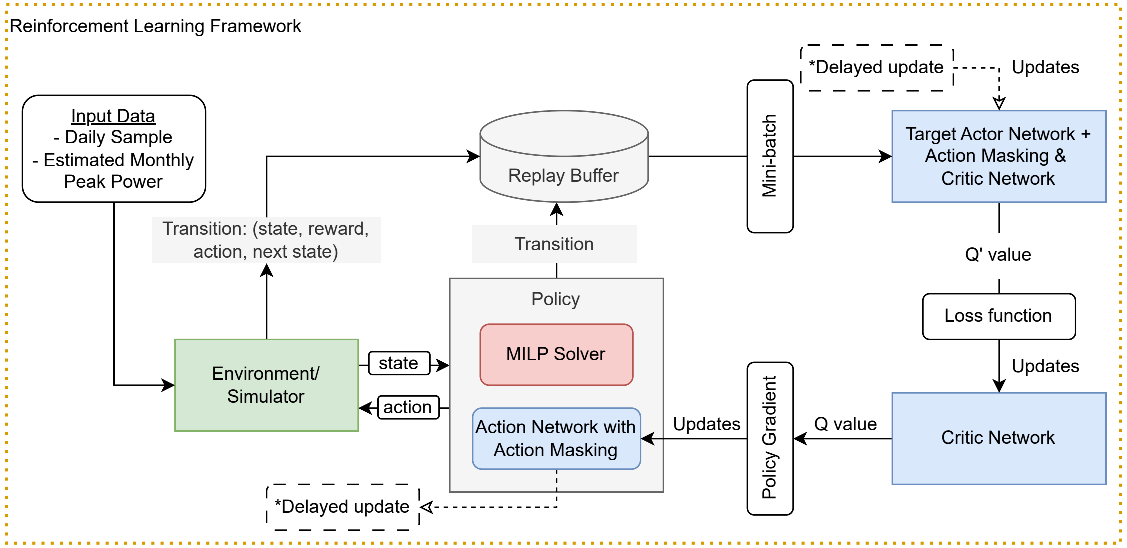
Opt-V2B-pen-TW:

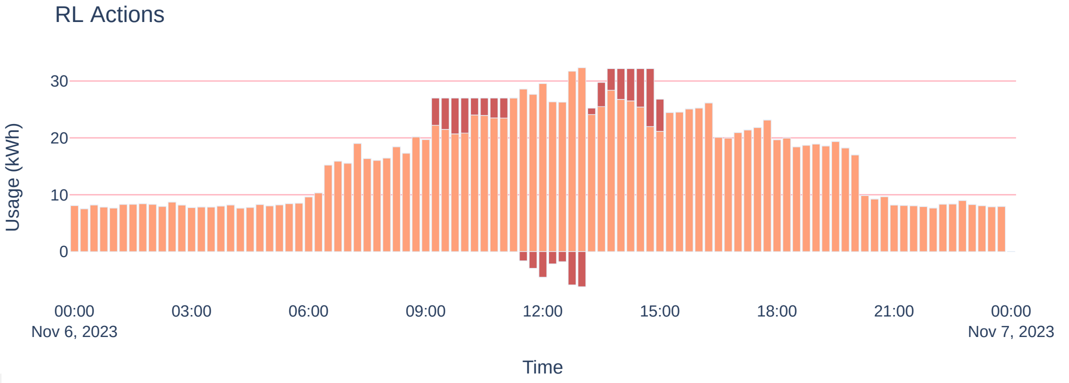
- **Allows flexible arrivals and any time departures** - Combines Opt-V2B-pen & Opt-V2B-TW

Solver	Demand ($)	Total Bill ($)	Uncharged cars
Opt-V2B	1046.29	1274.28	0

Actions under uncertainty: Reinforcement Learning model

- Uncertainty from the building, cars and grid.
- Not knowing the future makes it hard to give the best charger actions.
- Training a model requires representing the state of the system as numerical features that a model can learn. This can easily be an overwhelming amount of data to learn.
- The real-world state is reduced to 38 features, and the monthly billing period is reduced to daily chunks.





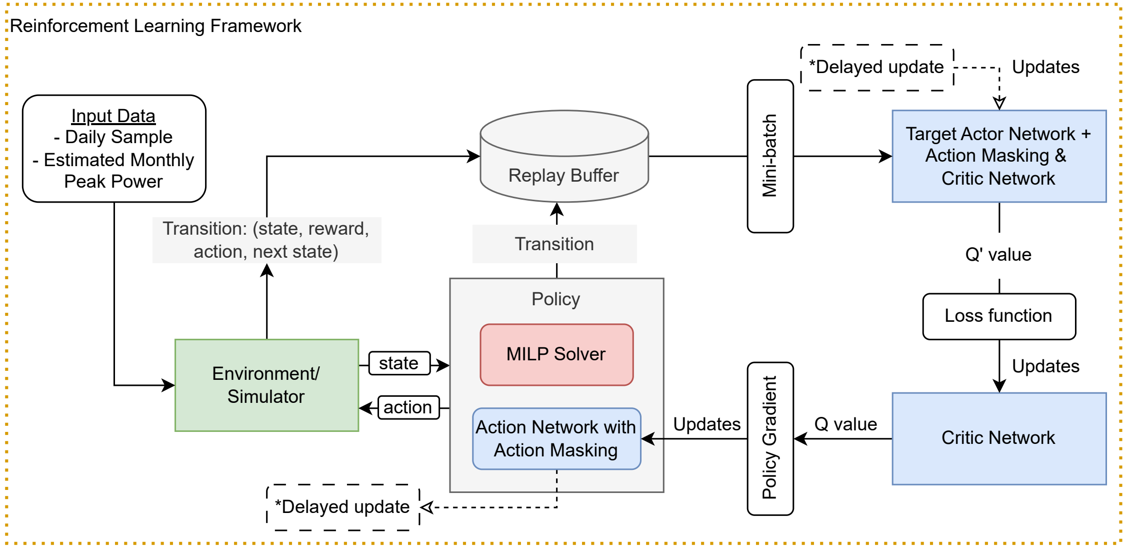
Solver	Demand ($)	Total Bill ($)	Uncharged cars
Opt-V2B	1046.29	1274.28	0
RL	1163.00	1391.00	0

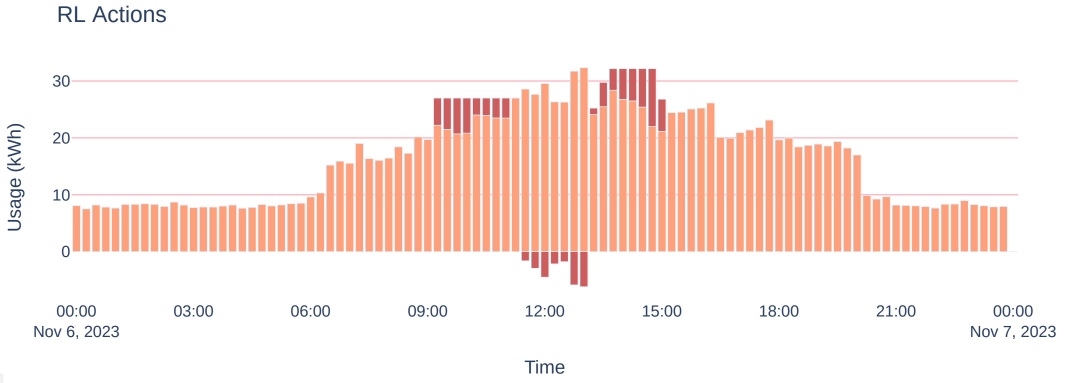
Actions under uncertainty: Reinforcement Learning model

- **DDPG**: RL training approach suitable for continuous action spaces, utilizing off-policy training, which facilitates better RL generalization.

- **Action Network with Action Masking**: We mask raw actions to feasible actions to guide RL training to better solutions.

- **Policy Guidance:** Integrating optimal actions from the MILP solution into the replay buffer to improve local optima.





Solver	Demand ($)	Total Bill ($)	Uncharged cars
Opt-V2B	1046.29	1274.28	0
RL	1163.00	1391.00	0

Reinforcement Learning Policy – required for Scale up

Data : Real-world data from Nissan's Santa Clara laboratory from May 2023 to January 2024.

Downsampling: We used k-means clustering on the 1,000 samples to select 60 for training and 50 for testing, ensuring exclusivity

Estimated Peak Power: We estimated this value based on the MILP solutions on training data.

Splitting Monthly Samples: We divide the monthly sample into daily episodes, focusing on weekdays to capture daily variations.

RL Training: We trained 9 models for all months, utilizing hyperparameter tuning through a k-fold cross-validation approach.

Comparison of RL and Baselines on Monthly Total Bill

Policy	MAY	JUN	JULY	AUG	SEP	OCT	NOV	DEC	JAN	Total
MILP	6201.1±50	6713.3±61	7371.0±40	9308.9±51	7231.0±36	7640.6±66	6625.9±42	6079.8±54	6495.1±55	64580.3
RL	**6222.6±26**	6857.1±122	**7392.2±51**	**9363.3±81**	**7243.0±24**	**7696.3±71**	**6654.9±61**	**6243.7±158**	**6635.0±80**	**64551.0**
CF-LLF	6245.9±32	**6843.4±42**	7396.8±26	9435.8±47	7284.1±41	7742.1±48	6675.9±32	6261.8±99	6646.3±81	64825.5
CF-EDF	6247.6±34	6849.6±48	7399.0±28	9436.1±47	7289.5±48	7747.6±49	6676.3±31	6276.6±87	6639.9±69	64806.2
T-LLF	6310.7±66	6920.0±75	7432.6±34	9537.5±52	7326.9±48	7800.1±48	6796.9±46	6344.5±132	6670.3±79	65045.5
T-EDF	6326.6±58	6920.0±56	7455.4±34	9543.0±54	7364.5±48	7819.7±57	6809.7±42	6356.4±88	6673.2±60	65028.4
Trickle	6333.8±44	6955.6±46	7506.0±37	9570.8±53	7402.1±47	7844.1±60	6842.9±44	6393.1±60	6706.8±53	65388.0
FC	6308.7±50	6968.6±72	7537.3±83	9541.7±61	7403.6±81	7804.0±69	6813.0±70	6646.9±144	6706.4±77	65292.3

- **MILP**: Optimal solution.
- **Fast Charge (CF)**: Charges EVsLLF but rate until the required SoC is reached.
- **Trickle:** Uses a minimum *trickle charging rate* (*Energy needed to reach the required SoC divided by the remaining stay duration.*) to meet the required SoC before departure.
- **Trickle LLF (T-LLF)**: Allocates the *power gap** to EVs using the trickle rate, prioritized by laxity - *Departure time minus the minimum charging time to reach required SoC.*
- **Trickle EDF (T-EDF)**: Similar to T-LLF, prioritizing based on departure time.
- **Charge First LLF (CF-LLF)**: Overcharges EVs when the power gap* (Estimated peak power minus building load.) isn't fully used by charging all EVs at the trickle charging rate, and discharges when it's insufficient, prioritizing by laxity*.
- **Charge First EDF (CF-EDF)**: Similar to CF-LLF, but prioritizes based on departure time.

Ablation Study: Monthly Total Bills from May to July 2023

RL	RL\500	RL\C	RL\F	RL\E	RL\P	RL\A	Random\A
20471.9 ± 137	20494.8 ± 174	20511.6 ± 184	20594.1 ± 181	21130.2 ± 214	21157.0 ± 204	21273.7 ± 209	21627.3 ± 180

- **RL\500**: Trained with 500 samples.
- **RL\C**: Trained using 60 randomly selected samples from 1000 generated.
- **RL\F**: Utilized all 100 complete state features.
- **RL\E**: Monthly estimated peak power is set to 0.
- **RL\P**: Training without policy guidance.
- **RL\A**: Training without action masking
- **Random\A**: Actions randomly selected instead of using a trained actor network, followed by action masking.

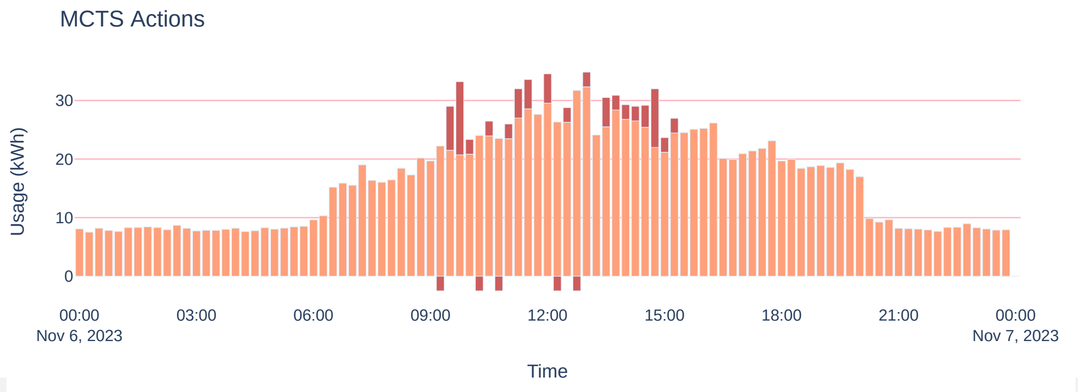
Results highlight the significance of downsampling, state definition, action masking, policy guidance, and estimated peak power.

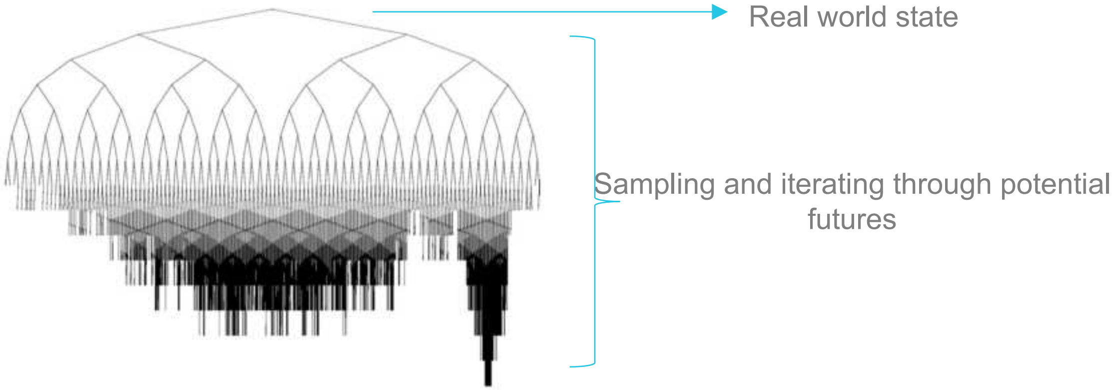
Ongoing Research: Monte Carlo Tree Search (MCTS)+ RL

- We can generate solutions for any given state by sampling potential future states and deciding on actions based on the potential rewards.

- Search algorithms can be very robust to uncertainty. By sampling potential futures as an approximation for the real-world future and iterating through it many times, it can arrive at a potentially great solution.
- This is very similar to the alphazero algorithm which has only been applied to game boards until now.
- We are working on updating it for our scenarios.

- This leads to very long and impractical run times.
- Efforts to decrease run times include decentralization or progressive widening which can lead to sub-optimal solutions.





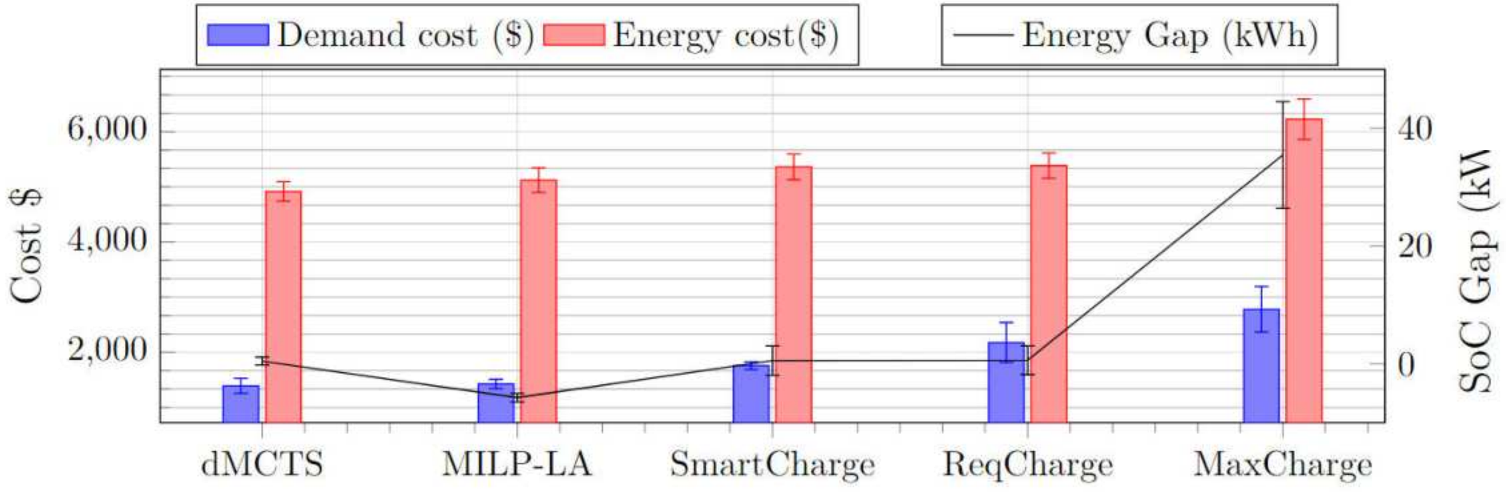
Adaptable – can update the underlying models when changes in the environment are detected

Challenge: scalability – convergence time can be a barrier for problems with large state-action spaces and limited decision-making time

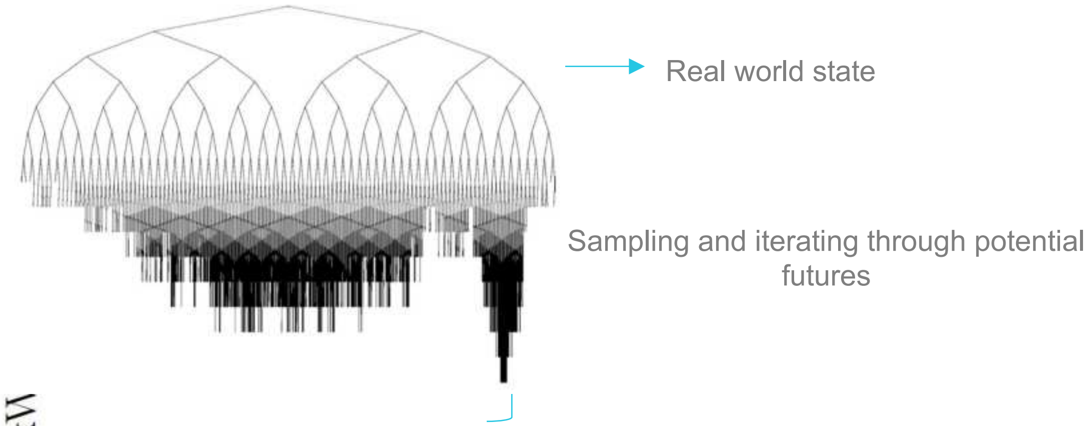
Ongoing Research: Monte Carlo Tree Search (MCTS)+ RL

■ Decentralization is key to manage the state complexity

- **MILP-LA** Applies optimal MILP action at each node of tree search
- **SmartCharge** minimizes peak usage by delaying charging, potentially missing required SoC.
- **ReqCharge** charges EVs quickly to the required SoC.
- **MaxCharge** charges EVs quickly to the maximum SoC limit.
- **dMCTS: A decentralized MCTS version implemented by the team.**



- This leads to very long and impractical run times.
- Efforts to decrease run times include decentralization or progressive widening which can lead to sub-optimal solutions.

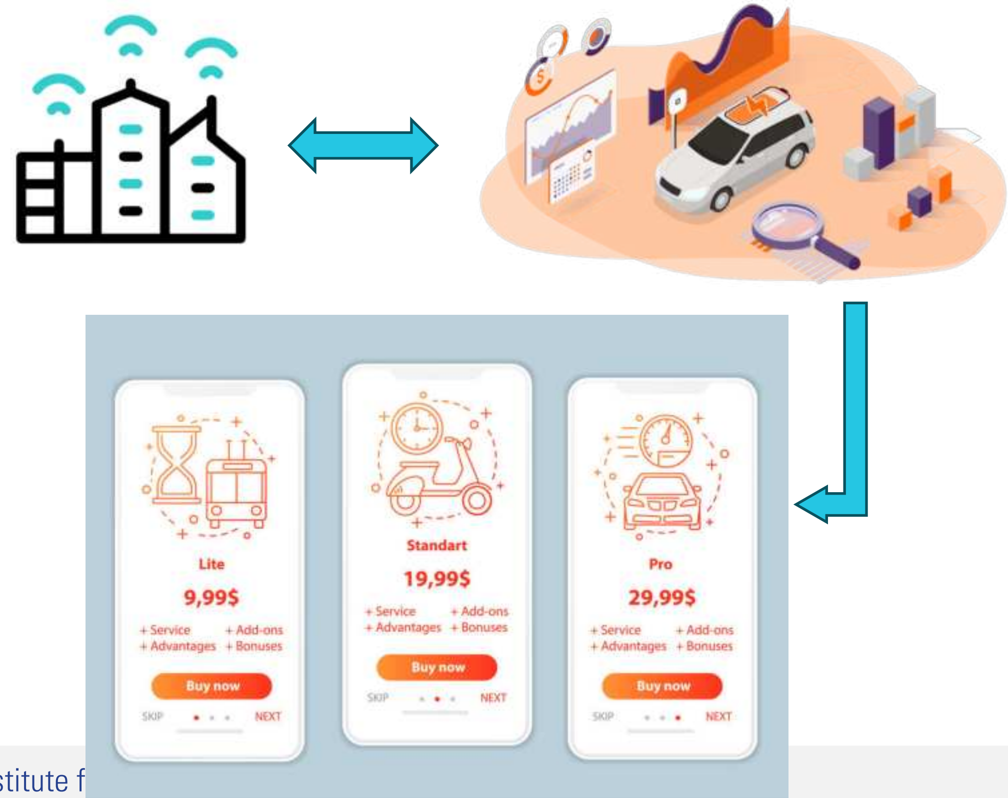


Looking to the future

■User Negotiations, Vehicle-to-Home, Vehicle-to-Grid

- Being able to provide the best charging actions is only valuable if users agree.
- A negotiation system needs to be in place to give options to users.

- Developing smart charging actions can help improve V2H and V2G interactions.

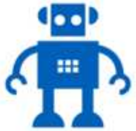


Acknowledgement: Project Team

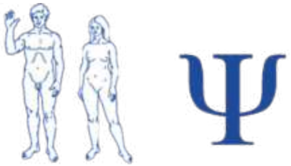
■Nissan Silicon Valley Team

Our Capabilities

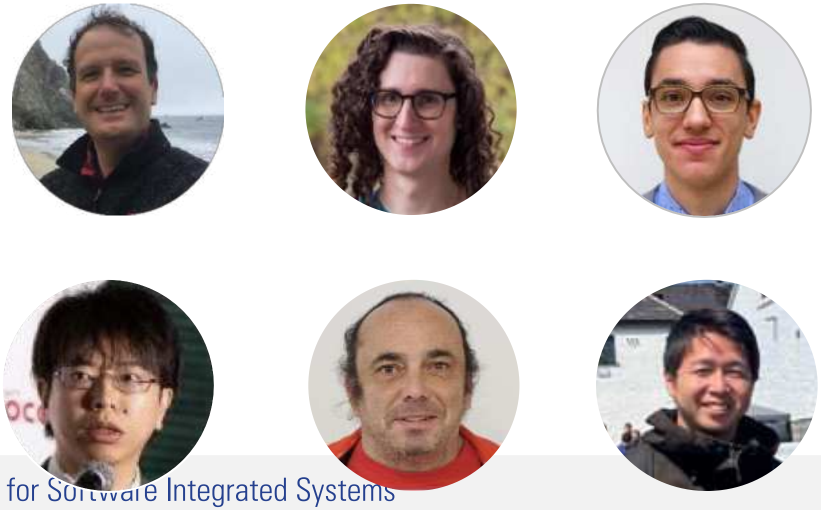

Computer Sciences


Robotics


Data Sciences


Social Sciences

Project Team



■Vanderbilt Team


AI for Cyber-Physical Systems


Resilient Design and Operation of Complex Cyber-Physical Systems


Decision making under uncertainty

Thank you for your attention.

Contact me for follow-ups

Abhishek Dubey, abhishek.Dubey@Vanderbilt.edu

Chancellor Faculty Fellow | Vanderbilt University
Director of Graduate Studies (MSc)| Computer Science
Associate Professor | Computer Science
Associate Professor | Electrical and Computer Engineering
Senior Research Scientist | Institute for Software Integrated Systems