

 dtec.bw

Zentrum für Digitalisierungs- und
Technologieforschung der Bundeswehr



HELMUT SCHMIDT
UNIVERSITÄT

Universität der Bundeswehr Hamburg

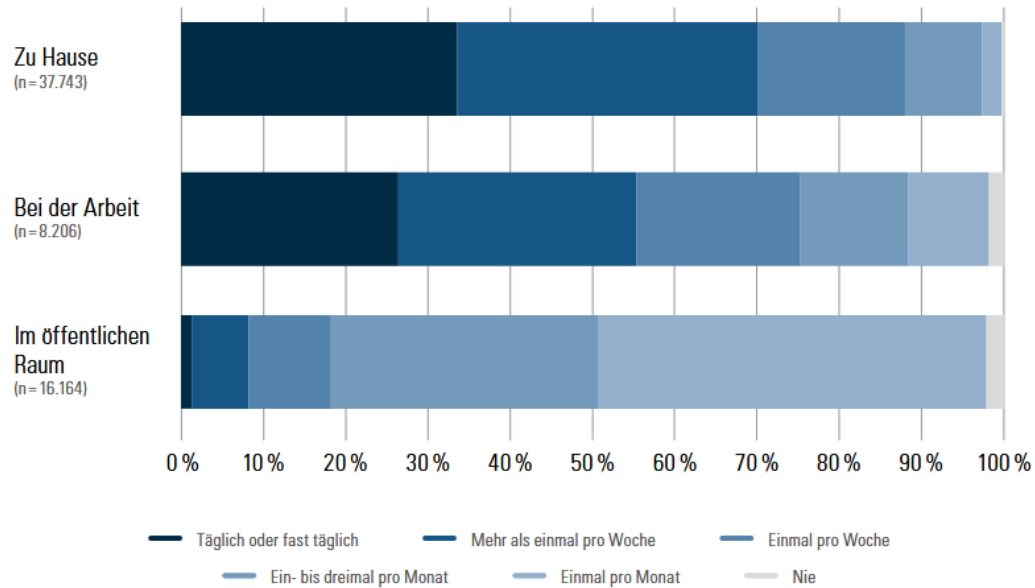


Netzlastmanagement via Ladesäulen mit Deep Reinforcement Learning unter Betrachtung verschiedener Beobachtungsräume

Dennis Salvador Versen

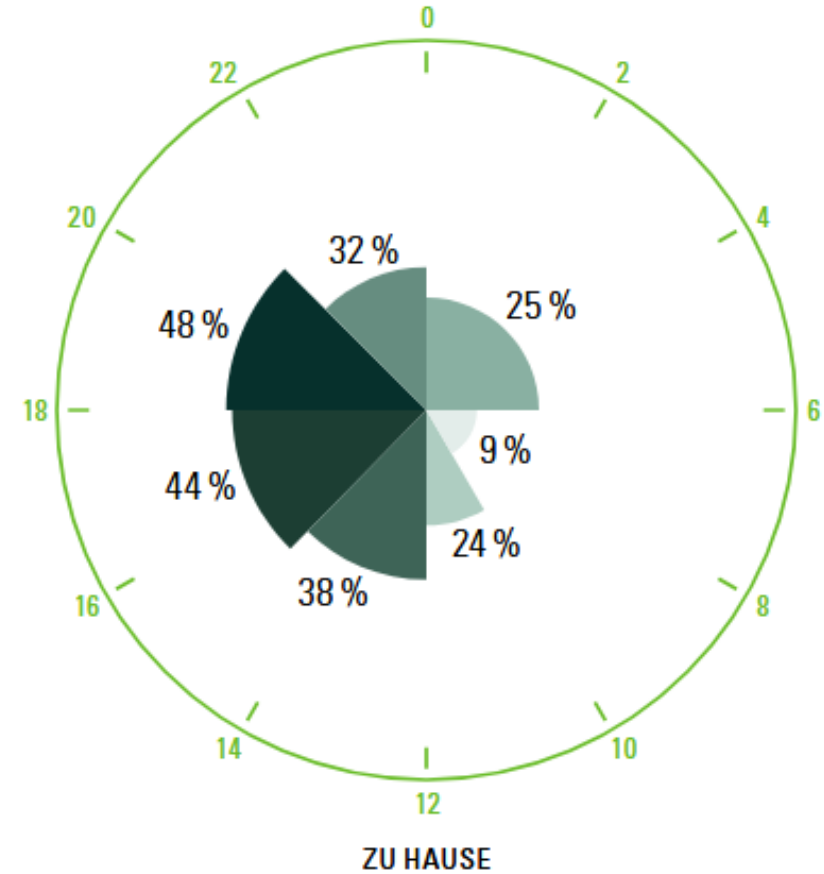
Yuzhuo Fu

Problemstellung



! „Befragte laden zu Hause am häufigsten.“

Angabe als Uhrzeit



Quelle: Nationale Leitstelle für Ladeinfrastruktur, „Einfach zu Hause laden“, 2024

Stand der Forschung

- Klassische Optimierung → hohe Modell- und Datenanforderung
- Viele Arbeiten nutzen synthetische Lastprofile → Gefahr der Mustererkennung zur Datengenerierung
- Klassisches Reinforcement Learning zeigt gute Ergebnisse hat aber Schwächen in komplexen Umgebungen mit stetigen Regelgrößen
- Model-Free-Deep-Reinforcement-Learning bietet hier eine modellfreie Alternative

Vorteile für Lastmanagement

- Kein physikalisches Modell erforderlich
- Datengetriebenes Lernen über Neuronale Netze
- Robustheit gegenüber Modellunsicherheiten
- Anpassung an stochastische Umgebungen
- Berücksichtigung langfristiger Regelstrategien

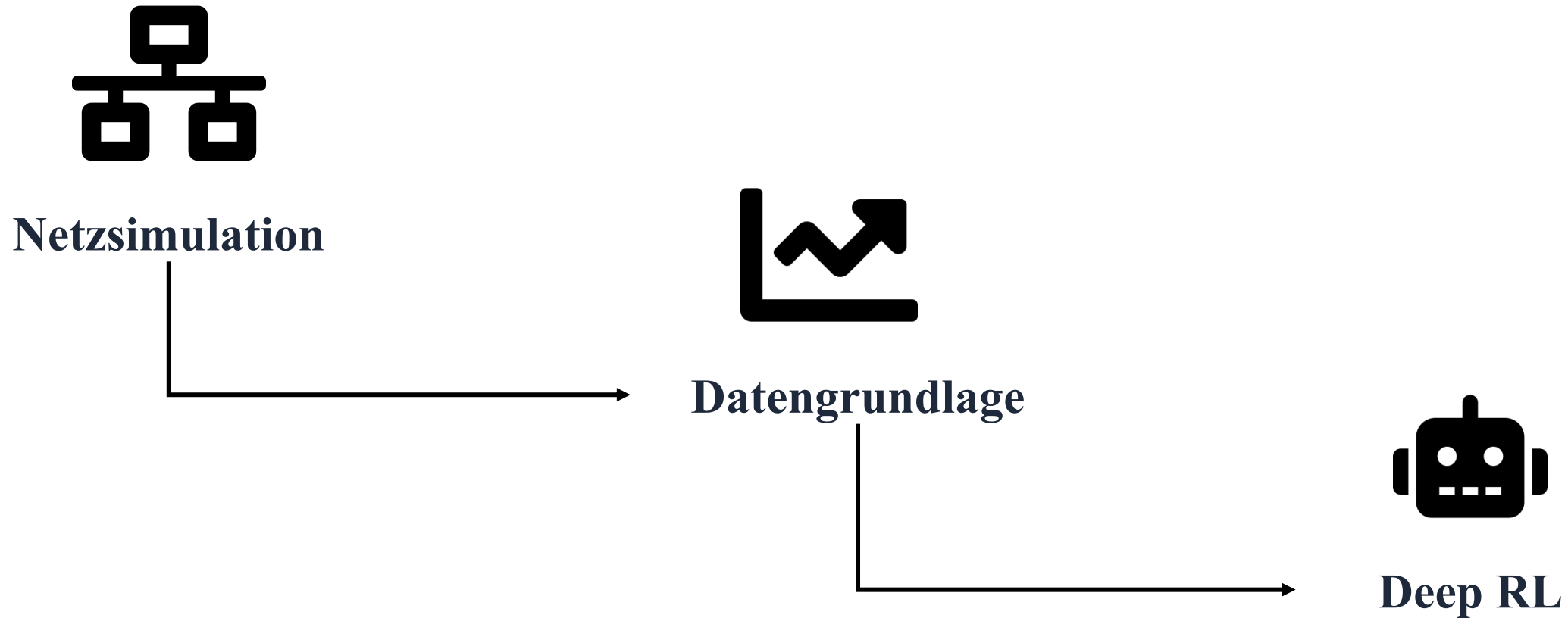
Zielsetzung & Forschungsfrage

- Präventive Steuerung von EV-Ladevorgängen durch Deep Reinforcement Learning (Deep RL) Agenten
- Untersuchung unterschiedlicher Informationsverfügbarkeiten
- Fokus auf reale Daten
- Vermeidung synthetischer, repetitiver Lastprofile

Zentrale Forschungsfrage:

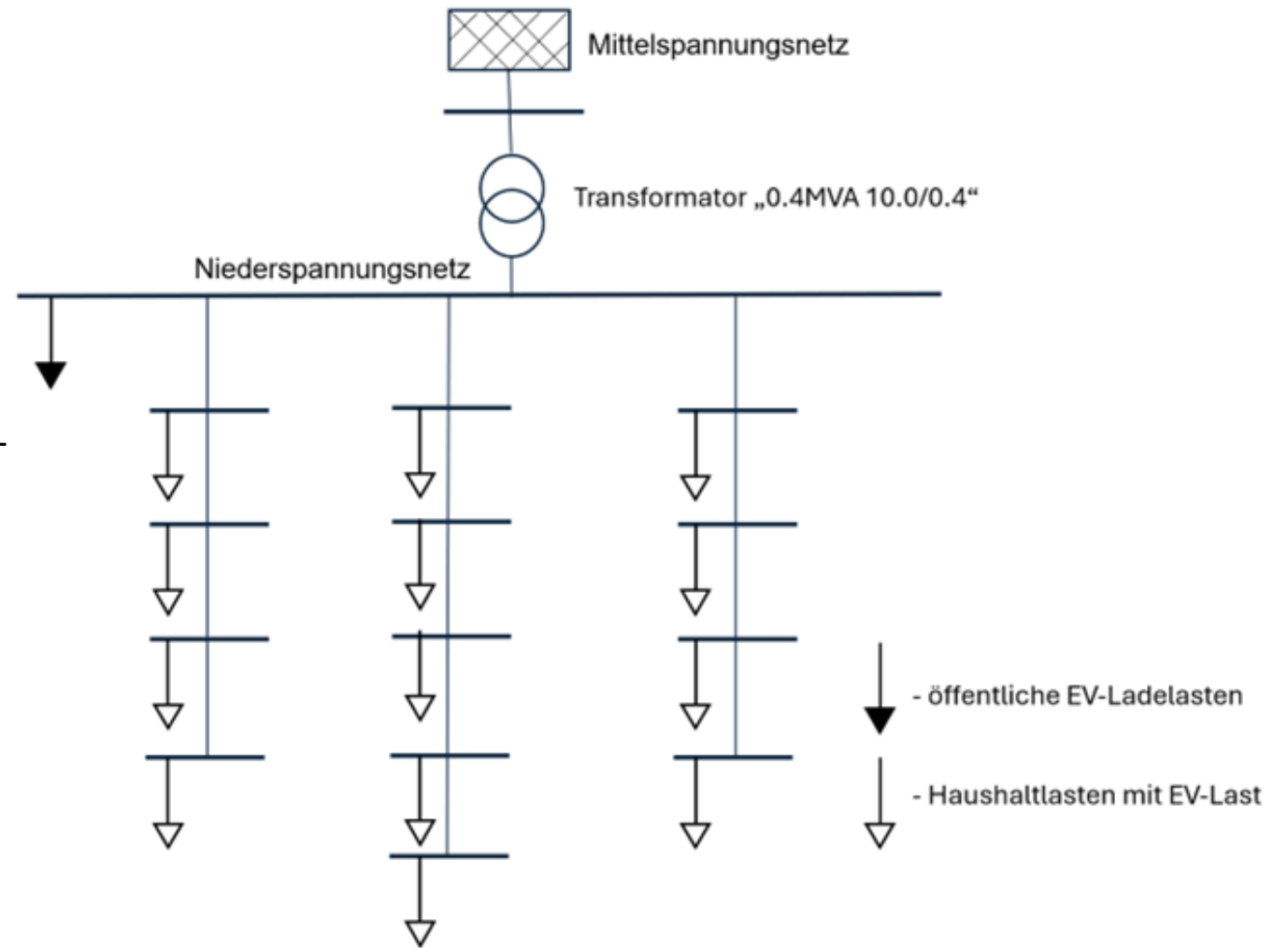
Kann ein prädiktives Echtzeit-Lastmanagement mit Reinforcement Learning in einer komplexen stochastisch dynamischen Umgebung mit realen Messdaten zur Netzstabilität beitragen, und welche Informationen sind hierzu relevant?

Methodik - Überblick



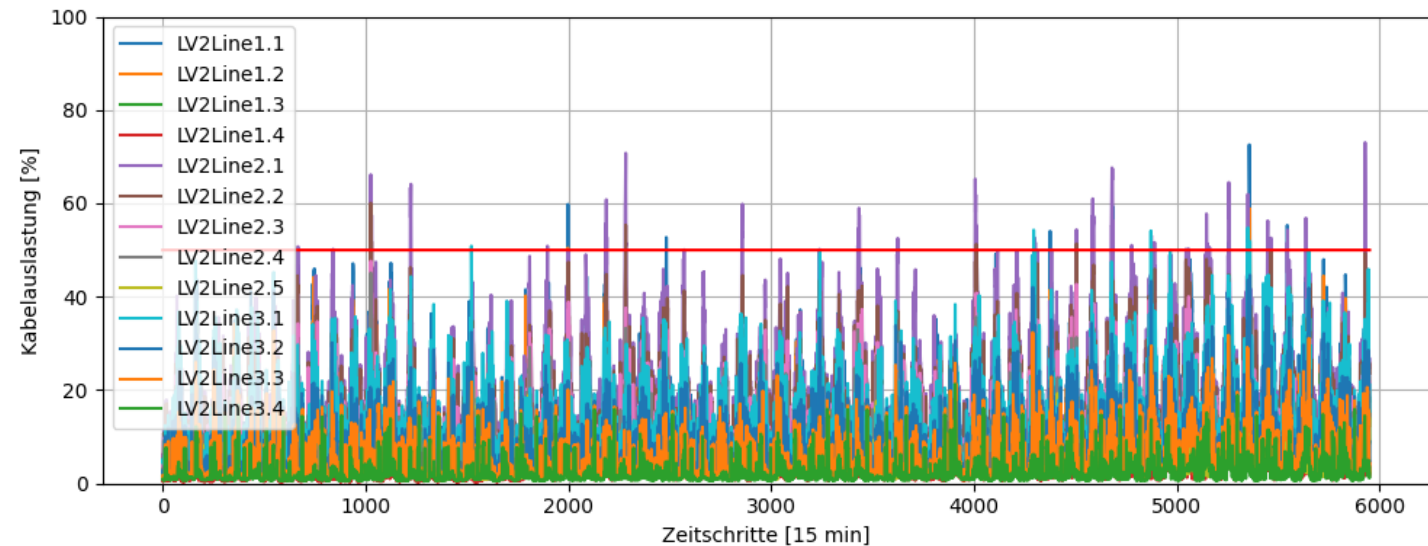
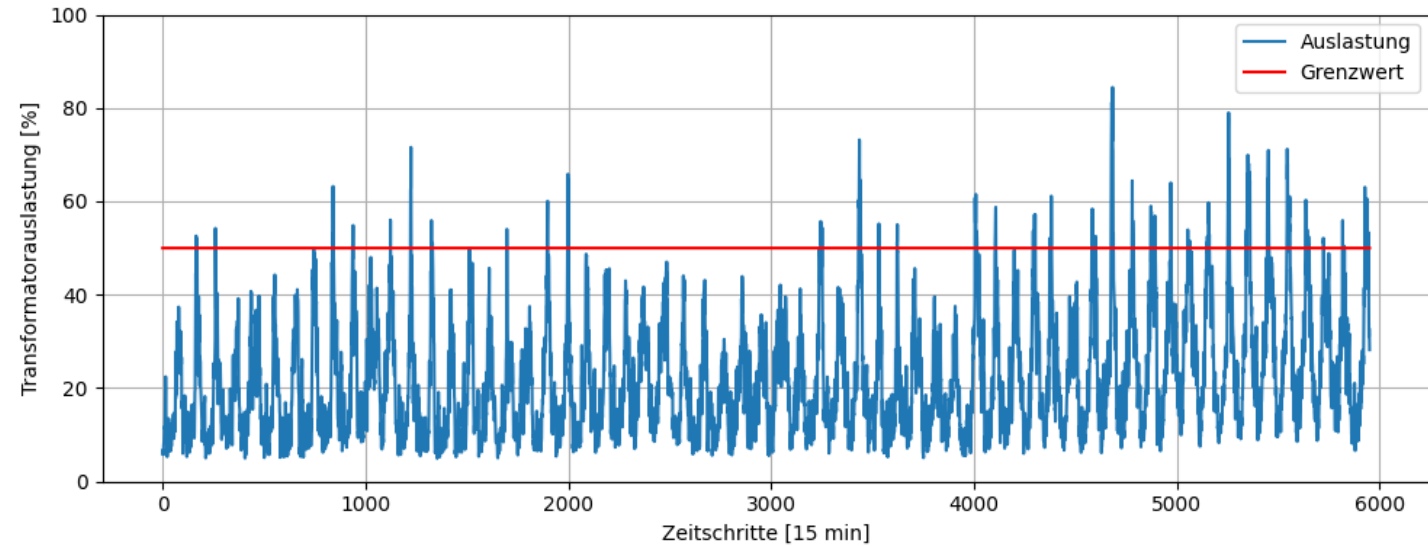
Systemmodellierung

- Simulation Niederspannungsnetz mit Pandapower
- 400-kVA-Ortsnetzstation
- Sammelschienen mit Haushaltslasten und hoher EV-Durchdringung
- Vektorisierung der Daten zur Verbesserung der Verarbeitungsgeschwindigkeit
- Lastflussberechnung pro 15-min-Zeitschritt

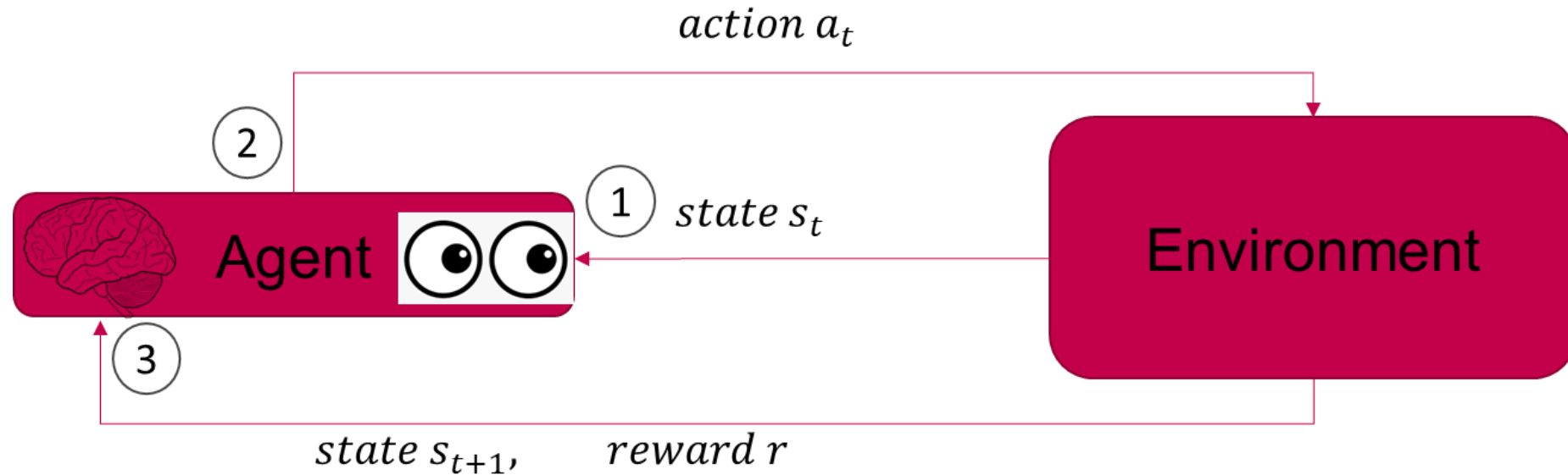


Datengrundlage

- Reale Haushalts- und EV-Daten (Pecan Street)
- Kombination aus Haushalts- und Ladeprofilen für den Zeitraum 15.04.-15.06.2015
- Aufteilung in Trainings- und Testdaten
- Grenzwertannahme: 50 % Auslastung von Betriebsmitteln



Reinforcement Learning



1. Interaktion mit der Umgebung
2. Erhalt von Feedback (reward) für Zustandsübergänge
3. Speichern der Zustandsübergänge und des Feedbacks
4. Training auf den gespeicherten Daten und Erlernen einer Strategie die das Feedback maximiert

Umgebung (Markovsches Entscheidungsproblem)

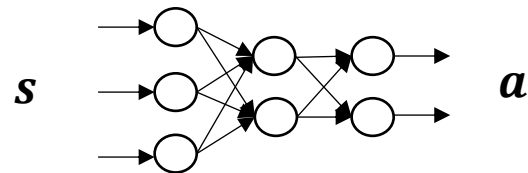
- Markovsches Tupel $\{S, A, P, R\}$
- Zustandsraum $s \in S \rightarrow$ Uhrzeit, Netz- und ggf. Kundendaten (Transformatorlast, Kabellast, EV-Daten etc.)
- Aktionsraum $a \in A \rightarrow$ kontinuierlicher Dimmfaktor $a \in \{0,1\}$
- Zustandsübergang $P(s'|s, a) \rightarrow$ Gegeben durch unsicheres Ladeverhalten
- Belohnungsfunktion $r = R(s, a)$

$$r = W \cdot \sum_i^N E_i - T \cdot \min(0, \theta - \theta_{ref}) - U \cdot \sum_l^L \min(0, \kappa_l - \kappa_{ref}) - V \cdot \Delta a$$

- θ - Transformatorauslastung
- κ - Kabelauslastung
- E_i - Geladene Energie von EV i
- W, T, U, V - Gewichte

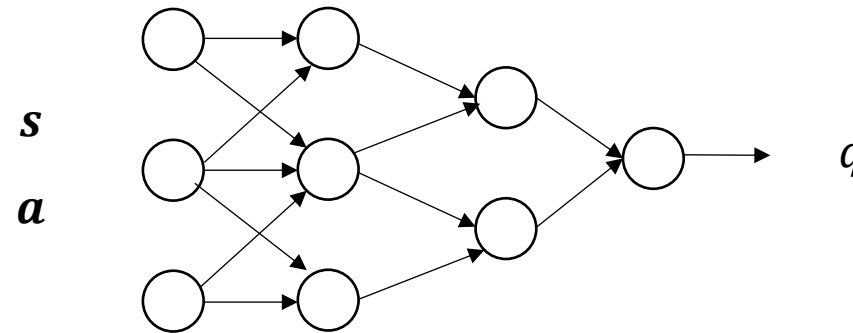
Deep Reinforcement Learning

Actor



$$a = \mu(s)$$

Critic



$$q = Q(s, a)$$

Training des Actor-Netzes:

$$\max_{\omega} Q(s_t, \mu_{\omega}(s_t))$$

Training des Critic-Netzes:

$$y = r + \gamma Q_{\phi}(s_{t+1}, \mu(s_{t+1})) \quad \gamma \in \{0,1\}$$

$$MSE: E(\phi) = (Q_{\phi}(s_t, a_t) - y)^2$$

$$\min_{\phi} E(\phi)$$

Deep RL Algorithmus im Vergleich

Eigenschaften	Deep Deterministic Policy Gradient (DDPG)	Soft Actor Critic (SAC)	Proximal Policy Optimization (PPO)
Typ	Off-Policy	Off-Policy	On-Policy
Policy	deterministisch	stochastisch	stochastisch
Sample Effizienz	hoch	hoch	niedrig
Stabilität	mittel	hoch	Sehr hoch
Konvergenzgeschwindigkeit	Schnell	Mittel	langsam

Szenarien (Beobachtungsräume)

Szenario 1:

- Nur netzseitige Messdaten, keine EV-Informationen
- $S_1 = \{\text{Zeitstempel}, \text{Transformatorlast}, \text{Kabellasten}\}$

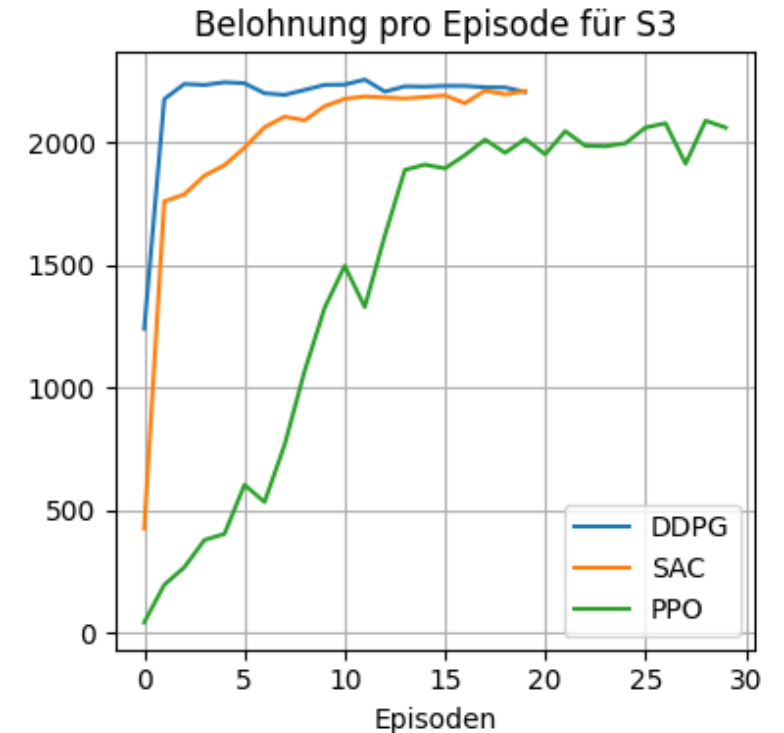
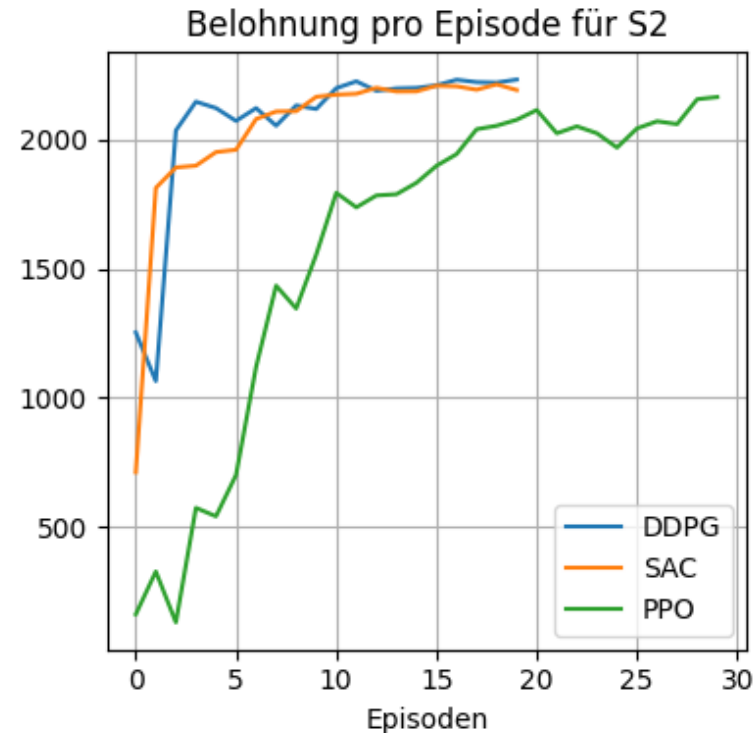
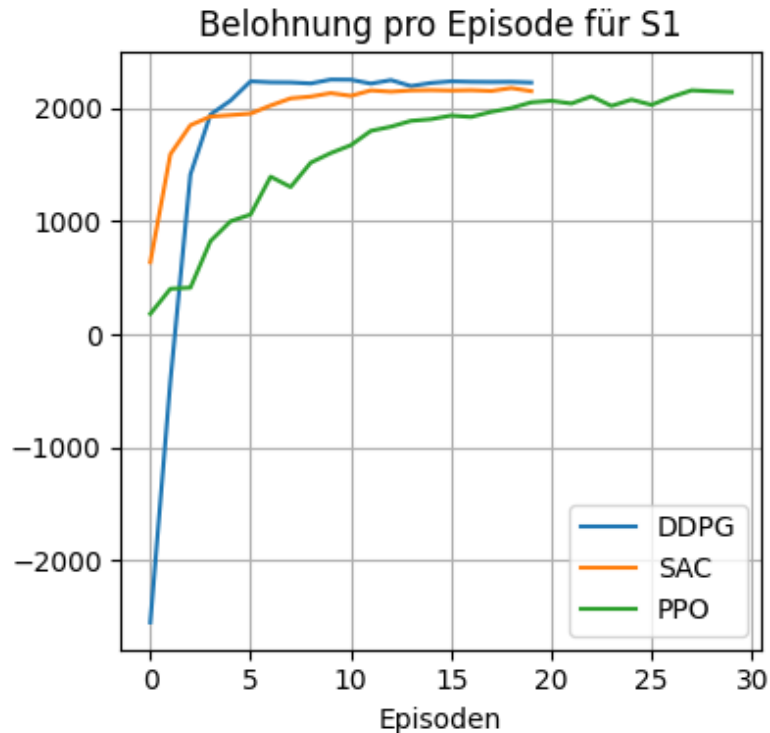
Szenario 2:

- Minimale Kundeninformation
- $S_2 = \{\text{Zeitstempel}, \text{Transformatorlast}, \text{Kabellasten}, \text{Anzahl der ladenden EVs in NSN – Gebiet}\}$

Szenario 3:

- Kumulierte Smart Meter Daten mit separatem Zähler für EV-Daten
- $S_3 = \{\text{Zeitstempel}, \text{Transformatorlast}, \text{Kabellasten}, \text{Ladeleistung}, \text{Haushaltslast}\}$

Trainingsergebnisse (Lernkurven)

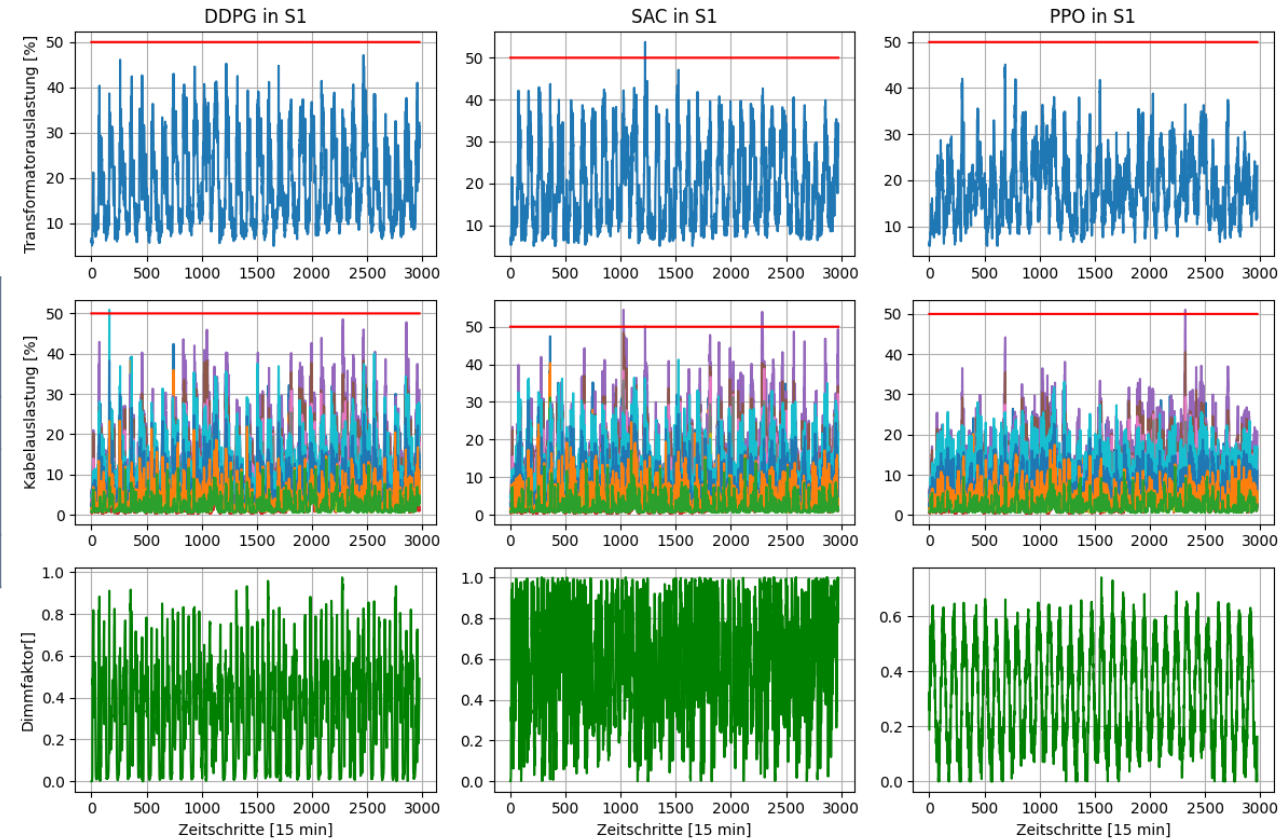


- DDPG und SAC konvergieren nach ~10 Episoden
- PPO benötigt ~30 Episoden (On-Policy Charakteristik)
- Alle Algorithmen zeigen erfolgreiche Strategiebildung
- DDPG und SAC erreichen höhere finale Scores

Evaluation: Szenario 1

Algorithmus	Trafo-Verletzungen	Kabel-Verletzungen	Max. Transformatorlast [%]	Max. Kabellast [%]	Ladeenergie [%]
DDPG	0	1 ($\approx 2,8\%$)	47,1 (34,2%)	50,9 (28,1%)	92,98
SAC	1 ($\approx 3,5\%$)*	5 ($\approx 13,9\%$)	53,8 (24,9%)	54,5 (23,0%)	96,29
PPO	0	1 ($\approx 2,8\%$)	45,1 (37%)	51,0 (28,0%)	91,62

*()-Klammern: Verhältnis zum unregulierten Zustand

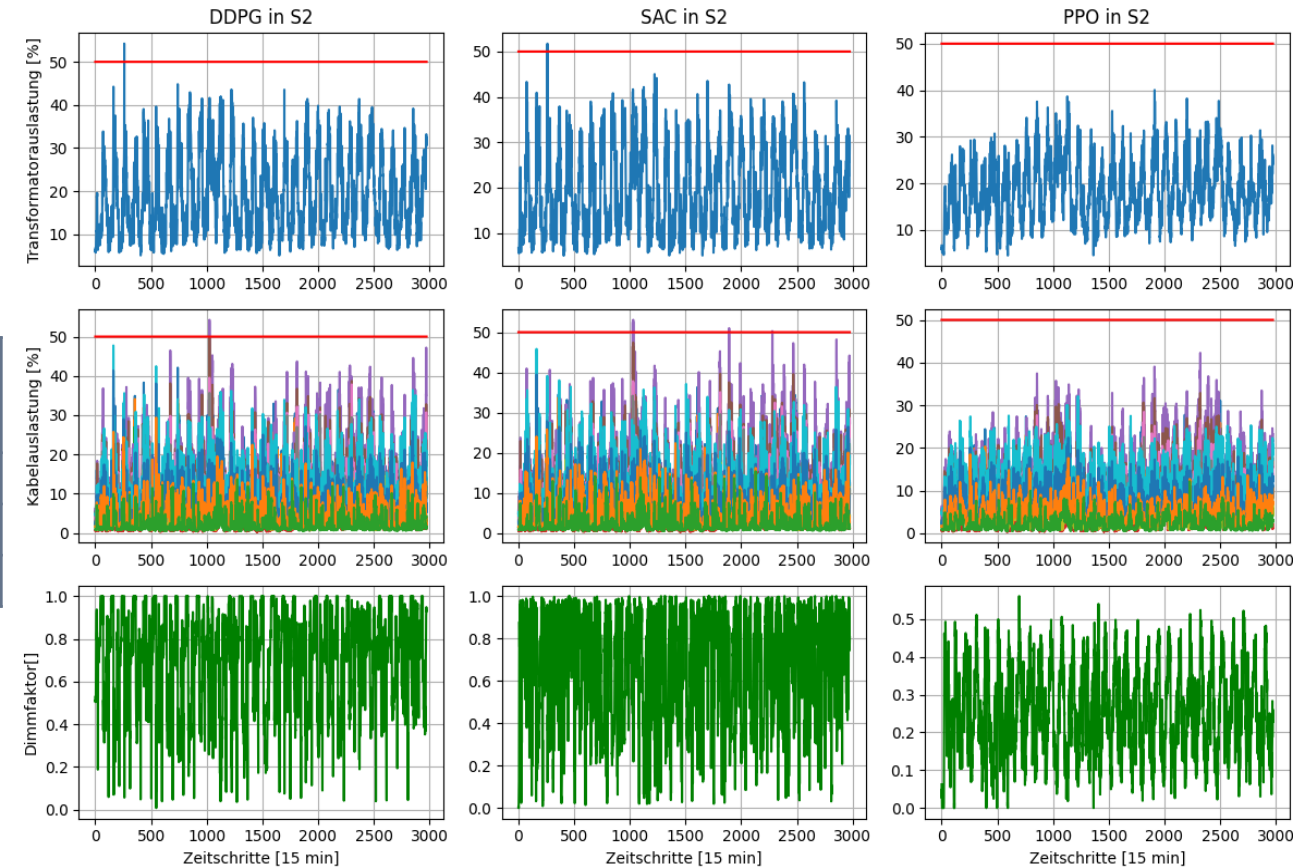


Darstellung der Transformatorauslastung, der Kabelauslastung und des Dimmfaktors im Test-Szenario 1

Evaluation: Szenario 2

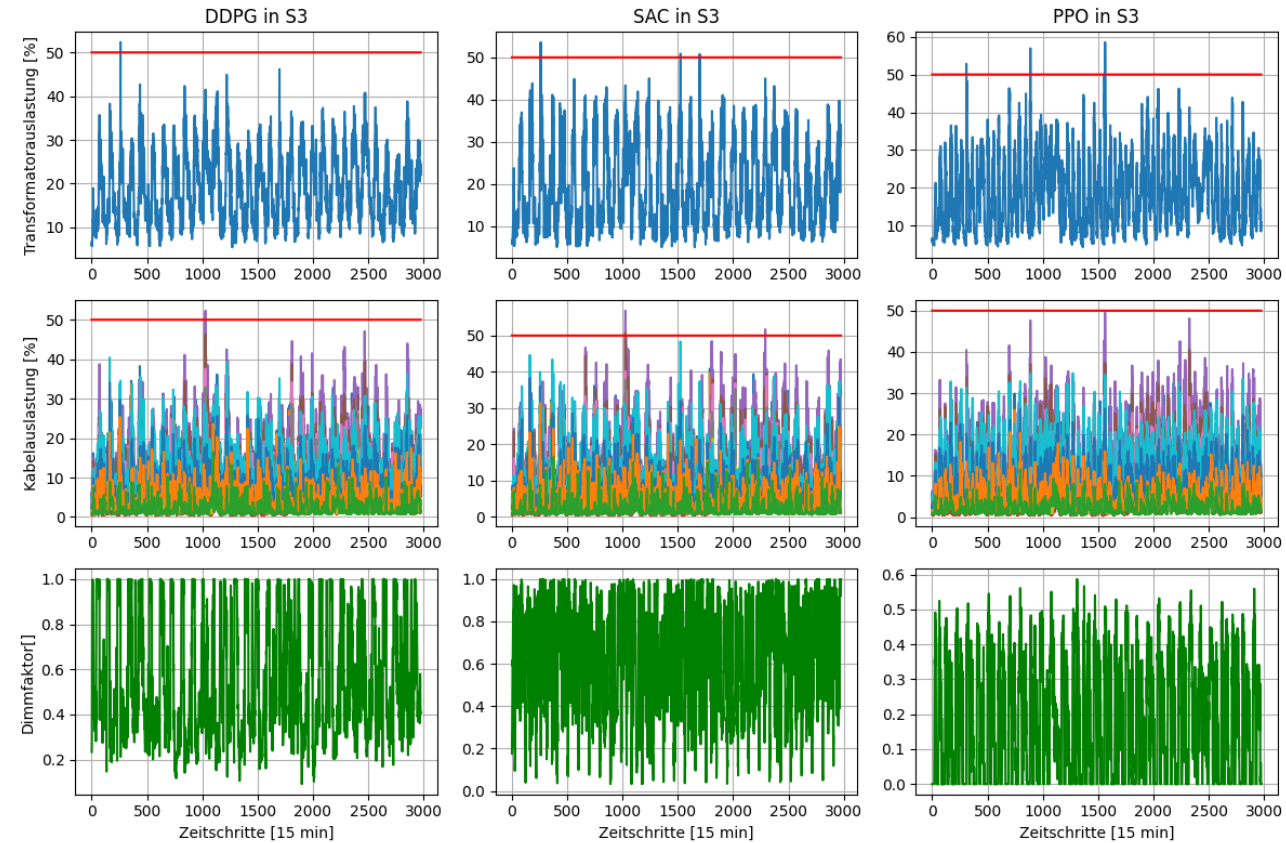
Algorithmus	Trafo-Verletzungen	Kabel-Verletzungen	Max. Transformatorlast [%]	Max. Kabellast [%]	Ladeenergie [%]
DDPG	1 (≈3,5%)	2 (≈ 5,6%)	54.2 (24,3%)	54,3 (23,3%)	96,44
SAC	1 (≈3,5%)	5 (≈13,9%)	51.7 (27,8%)	53,1 (25,0%)	96,72
PPO	0	0	40,1 (43,9%)	42,3 (40,3%)	90,94

*)-Klammern: Verhältnis zum unregelten Zustand



Darstellung der Transformatorauslastung, der Kabellastung und des Dimmfaktors im Test-Szenario 2

Evaluation: Szenario 3



Darstellung der Transformatorauslastung, der Kabellastung und des Dimmfaktors im Test-Szenario 3

Algorithmus	Trafo-Verletzungen	Kabel-Verletzungen	Max. Transformatorlast [%]	Max. Kabellast [%]	Ladeenergie [%]
DDPG	1 (≈3,5%)	2 (≈ 5,6%)	52,3 (27,0%)	52,3 (26,1%)	94,66
SAC	3 (≈10,7%)	4 (≈11,1%)	53,6 (25,1%)	56,9 (19,6%)	97,29
PPO	4 (≈14,3%)	0	58,5 (18,3%)	49,5 (30,0%)	89,57

*()-Klammern: Verhältnis zum unregulierten Zustand

Fazit

Vorteile

- Modellfreies Lernen
- Funktioniert effektiv mit realen Messdaten aus komplexen stochastischen Umgebungen
- Minimale Informationen (nur netzseitige Daten) zeigen gutes Ergebnis (Datenschutzkonform)
- Generalisierung auf unbekannte Daten
- SAC zeigt bestes Ergebnis
- Alternative Begleittechnologie zum Netzausbau

Limitationen

- Stochastische Ergebnisse (abhängig von Initialisierung, Hyperparameter)
- Netzkapazität nicht vollständig ausgeschöpft
- Aktuelle reale Trainingsdaten sind schwer zu erhalten
- Bidirektionales Laden nicht berücksichtigt
- Keine dynamischen Strompreise

SAC	Beobachtungsraum	Ladeenergie [%]	Grenzwertverletzungen
S1	Zeit, θ , κ	96,29	6
S2	Zeit, θ , κ , N	96,72	6
S3	Zeit, θ , κ , ΣP_{CP} , ΣP_{HH}	97,29	7

Ausblick & Weiterführende Forschung

Potenzielle Erweiterung

- Integration bidirektionalen Ladens (Vehicle-to-Grid)
- Berücksichtigung dynamischer Strompreise
- Einbindung lokaler dezentraler Erzeuger (PV)
- Optimierung der Hyperparameter für höhere Netzauslastung
- Erweiterung des Aktionsraums (individuelle Steuerung)
- Multi-Agent Reinforcement Learning

Praktische Umsetzung

- Feldtests in realen Niederspannungsnetzen
- Kontinuierliches Lernen mit aktuellen Daten

Referenzen

- [1] „Umwelt Bundesamt,“ 15 04 2025. [Online]. Available: <https://www.umweltbundesamt.de/themen/klima-energie/internationale-klimapolitik/uebereinkommen-von-paris/>.
- [2] „European Parliament,“ 08 08 2023. [Online]. Available: <https://www.europarl.europa.eu/topics/en/article/20180920STO14027/reducing-car-emissions-new-co2-targets-for-cars-and-vans-explained>.
- [3] B. d. Justiz, Gesetz über die Elektrizitäts- und Gasversorgung (Energiewirtschaftsgesetz - EnWG), EnWG: https://www.gesetze-im-internet.de/enwg_2005/_14a.html, 2005 (Jahr der Bekanntmachung).
- [4] S. Schey, D. Scofield und J. Smart, „A First Look at the Impact of Electric Vehicle Charging on the Electric Grid in The EV Project,“ in EVS26 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium, Los Angeles, California, 2012.
- [5] Nationale Leitstelle, „Einfach zu Hause laden – Studie zum Ladeverhalten von Privatpersonen mit Elektrofahrzeug und eigener Wallbox,“ 2025.
- [6] V. F. Netztechnik/Netzbetrieb, „Ein Meilenstein auf dem Weg zu intelligenten Verteilnetzen,“ Bundesnetzagentur, Berlin, 2023.
- [7] „Bundesnetzagentur,“ [Online]. Available: <https://www.bundesnetzagentur.de/DE/Vportal/Energie/SteuerbareVBE/artikel.html>.
- [8] Y. Fu und D. Vesen, „Electric Vehicle Charging Management for Avoiding Transformer Congestion Using Policy-based Reinforcement Learning,“ in IEEE, 2023.
- [9] D. Vesen, „Entwicklung eines netzseitigen Lastmanagements auf Basis intelligenter nichtlinearer Systemidentifikation zur Vermeidung von Transformator-Lastspitzen in einem Niederspannungsnetz,“ in TKB, 2025.
- [10] J. Zhang, X. Zhan, T. Li, L. Jiang, J. Yang, Y. Zhang und X. Diao, „A Convex Optimization Algorithm for Electricity Pricing of Charging Stations,“ MDPI, Bd. 12, Nr. 10 - Recent Advances in Nonsmooth Optimization and Analysis, 2019.
- [11] H. Ameer, Y. Wang, X. Fan und Z. Chen, „Hybrid optimization of EV charging station placement and pricing using Bender’s decomposition and NSGA-II algorithm,“ Science Direkt , Bd. 397, Nr. Applied Science, 2025.
- [12] H. Sun, P. Yuan, Z. Sun, S. Hu, F. Peng und W. Zhou, „Distribution Network Congestion Dispatch Considering Time-Spatial Diversion of Electric Vehicles Charging,“ Energies (MDPI), Bd. 11, Nr. 10, pp. 1-17, 2018.
- [13] D. Liu, P. Zeng, S. Cui und C. Song, „Deep Reinforcement Learning for Charging Scheduling of Electric Vehicles Considering Distribution Network Voltage Stability,“ MDPI, Bd. 23, Nr. Optimal Planning, Integration and Control of Smart Grids and Microgrids Systems, p. 1618, 2022.
- [14] S. Sykiotis, C. Menos-Aikateriniadis, A. Doulamis, N. Doulamis und P. S. Georgilakis, „A self-sustained EV charging framework with N-step deep reinforcement learning,“ Elsevier, Bd. 35, Nr. Sustainable Energy, Grids and Networks, p. 101124, 2023.

Referenzen

- [15] B. d. J. u. Verbraucherschutz, Gesetz über den Messstellenbetrieb und die Datenkommunikation in intelligenten Energienetzen I (Messstellenbetriebsgesetz - MsbG), Regelungen zur Datenkommunikation in intelligenten Energienetzen §§ 49 bis 52 .
- [16] F. L. D. Silva, C. E. H. Nishida, D. M. Roijers und A. H. R. Costa, „Coordination of Electric Vehicle Charging Through Multiagent Reinforcement Learning,“ IEEE, Bd. 11, Nr. 3, pp. 2347 - 2356, 2019.
- [17] F. L. D. Silva, C. E. H. Nishida, D. M. Roijers und A. H. R. Costa, „Coordination of Electric Vehicle Charging Through Multiagent Reinforcement Learning,“ IEEE Transactions on Smart Grid, Bd. 11, Nr. 3, pp. 2347 - 2356, 2020.
- [18] A. Poddubnyy, P. Nguyen und H. Sloomweg, „Online EV charging controlled by reinforcement learning with experience replay,“ Elsevier, Bd. 36, Nr. Sustainable Energy, Grids and Networks, p. 101162, 2023.
- [19] pandapower, „pandapower.readthedocs.io,“ 2016-2024 . [Online]. Available: <https://pandapower.readthedocs.io/en/latest/>. [Zugriff am 2024].
- [20] „PECAN STREET,“ [Online]. Available: <https://www.pecanstreet.org/dataport/>.
- [21] M. Lapan, Deep Reinforcement Learning, mitp, 2020.
- [22] T. P. Lillicrap, J. J. Hunt, A. Pritzel und N. Heess, „CONTINUOUS CONTROL WITH DEEP REINFORCEMENT,“ in arXiv, London, 2016.
- [23] A. Z. P. A. S. L. Tuomas Haarnoja, „Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor,“ in arXiv, 2018.
- [24] J. Schulman, F. Wolski, P. Dhariwal, A. Radford und O. Klimov, „Proximal Policy Optimization Algorithms,“ in arXiv, 2017.
- [25] N. Andrenacci und M. P. Valentini, „A Literature Review on the Charging Behaviour of Private Electric Vehicles,“ MDPI, p. 29, 2023.
- [26] D. An, F. Cui und X. Kang, „Optimal scheduling for charging and discharging of electric vehicles based on deep reinforcement learning,“ Frontiers, Bd. 11, Nr. Sec. Energy Efficiency, 2023.
- [27] Verbraucherzentrale, „www.verbraucherzentrale.de,“ 25 06 2025. [Online]. Available: <https://www.verbraucherzentrale.de/wissen/energie/preise-tarife-anbieterwechsel/>.
- [28] S. Ayyadi, H. Bilil und M. Maaroufi, „Optimal charging of Electric Vehicles in residential area,“ Science Direkt, Bde. %1 von %2Sustainable Energy, Grids and Networks, Nr. 19, 2019.
- [29] R. Zhang, Z. Li, C. Wei, Y. Li, J. Yang und S. Su, „Optimization and Solution Method for Electric Vehicle Charging and Discharging Load,“ in IEEE 4th International Electrical and Energy Conference, Wuhan, China , 2021.