

DEEP REINFORCEMENT LEARNING ALS METHODE ZUR AUTONOMEN STEUERUNG VON NIEDERSPANNUNGSNETZEN

Lars QUAKERNACK*¹, Michael KELKER, Jens HAUBROCK^{Fehler! Textmarke nicht definiert.}

Einleitung

Das Bestreben ein emissionsarmes elektrisches Netz in Deutschland zu erreichen birgt für Verteilnetzbetreiber Herausforderungen. Dies ist unter anderem der Ausbau volatiler und dezentraler erneuerbarer Energieanlagen (DEA), wie zum Beispiel Photovoltaik und der stetige Ausbau an Elektromobilität mit teils hohen Ladeleistungen in der Niederspannung. Damit es zu keinen Einbußen in der Netzstabilität kommt, müssen entweder Steuerungssysteme für die Verbraucher und Erzeuger implementiert werden oder das elektrische Netz massiv ausgebaut werden. Da letzteres viel Geld kostet, forschen Wissenschaftler an Steuerungssystemen. Ein vielversprechender Ansatz ist die Steuerung mittels künstlicher Intelligenzen (KI) - genauer mittels Reinforcement Learning (RL). In diesem Beitrag wird RL als Methode für Niederspannungsnetze analysiert. Hierzu wird die Methode theoretisch erläutert und die Adaption auf das Niederspannungsnetz mit dem aktuellen Stand der Forschung dargelegt.

Methodik des Reinforcement Learning

RL gehört zu den Methoden des Maschinellen Lernens. Das zu Deutsch bestärkende Lernen sieht vor mittels Belohnungsfunktionen (Reward) für den Zustand der Umgebung im aktuellen Zeitschritt Agenten zu trainieren die beste Aktion zu wählen, um ein definiertes Optimum zu erreichen.

Für den Lernprozess werden im Zeitschritt t_i für jede Aktion a_t und Zustand s_t , Aktions/Zustands Paare als Q-value gebildet und in sogenannte Q-tables gespeichert. Auf Basis des Zustandes s_t wird der Reward r_t berechnet. Die Q-values werden im Training nach der Bellman Gleichung (1) in der Tabelle iterativ aktualisiert. Der höchste Q-value für einen Zustand bestimmt die Aktion. Dieses Verfahren wird Q-Learning genannt [4].

$$Q_{\text{new}}(s_t, a_t) \leftarrow Q_{\text{old}}(s_t, a_t) + \alpha \times [r_t + \gamma \times \max(Q(s_{t+1}, a_t) - Q_{\text{old}}(s_t, a_t))] \quad (1)$$

In Gleichung (1) beschreibt die Learning Rate α , inwiefern neue Erfahrungen des Agenten alte überlagern. Der Discountfaktor γ bestimmt die Wichtigkeit zukünftiger Rewards [4,5].

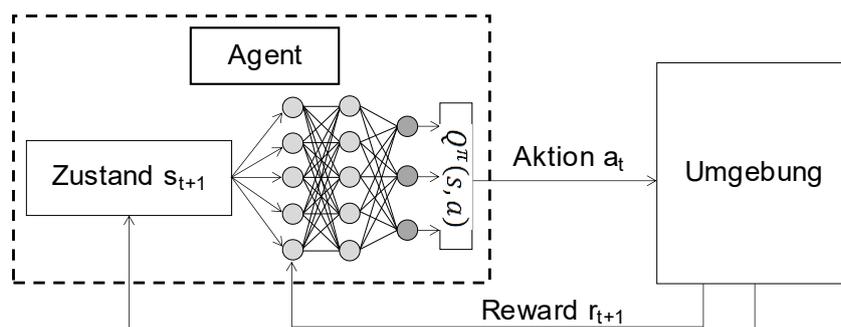


Abbildung 1: Schematische Darstellung von DQN

In Abbildung 1 ist schematische die Struktur von RL bzw. von dem für diese Problematik meist verwendete Deep Q-Learning (DQN) abgebildet. [1,2,3]. Der Unterschied zu dem einfachen Q-learning ist die Nutzung neuronaler Netze zur Vorhersage der Q-values. Dies wird gemacht, da die Q-Tables bei komplexeren Problemen zu Speicherplatzproblemen führen, da jede Aktion/State Kombination festgehalten werden muss. DQNs werden typischerweise entweder als Single Agent Reinforcement Learning (SARL) oder Multi Agenten Reinforcement Learning verwendet (MARL) [4,5].

¹ Fachhochschule Bielefeld Institut für Technische Energie-System (ITES), Interaktion 1 33619 Bielefeld Deutschland, +49.521.106-70341, lars.quakernack@fh-bielefeld.de, <https://www.fh-bielefeld.de/iium/forschung/agnes>

RL in Niederspannungsnetzen

RL dient in den Niederspannungsnetzen als Controller, um Vermeidung von Überlastungen von Betriebsmitteln, Ausgleich von DEA durch Batteriespeicher oder die Einhaltung von Spannungsgrenzen zu erreichen. Besonders D. Didden [1], J. Duan [2] und M. Kelker [3] haben in Ihren Arbeiten diese Art von Steuerung untersucht. In Tabelle 1 werden die Parameter der Arbeiten dargestellt:

Tabelle 1: Verschiedene DQN Parameter

	D. Didden [1]	J. Duan [2]	M. Kelker [3]
Umgebung	Nicht spezifiziert, verschiedene	200-bus System	12 Knoten, 2 Stränge Niederspannungsnetz
Zustand	Kalenderdaten, SoC Batterie, Nutzleistungsprognose	Spannungsdaten	Strom- und Spannungsdaten, Zeit, alte Aktionen
Reward	$R_{voltage}$, R_{line} , R_{trafo} , R_{losses}	$R_{voltage}$	$R_{voltage}$, R_{trafo}
Agent	SARL, MARL	SARL	MARL
Aktionen	Batterieleistung	Einstellwerte von Generatoren	Ladeleistung, Batterieleistung

D. Didden untersucht in seiner Arbeit den Einsatz netzdienlicher Batteriespeicher zur Vermeidung von Überlast an Transformator und Leitungen und Über- und Unterspannung. J. Duan nutzt die Erregung der Generatoren, um die Spannung im Netz zu kontrollieren. M. Kelker benutzt zusätzlich zu den Batteriespeichern, Elektrofahrzeuge als Steuerungselemente, um Überlast zu vermeiden. Die Ergebnisse von M. Kelker sehen eine Reduktion der Transformatorleistung von 24.4% und Erhöhung des Anteils an EE von 10 % durch die Steuerung vor. Bei J. Duan braucht der Agent in 99.92% der Fälle nur eine Iteration, um die richtige Aktion für die Spannungshaltung zu finden. Die Ergebnisse bei D. Didden zeigen, dass z.B. die Spannungsbandverletzungen durch DQN deutlich reduziert werden können. Die betrachteten Autoren kommen zu dem Ergebnis, dass RL ein großes Potential für die Netzsteuerung hat [1,2,3]. RL-Modelle lernen auf Basis ihrer Umgebungen. Hierbei sind die RL Umgebungen in [1-3] Simulationen des elektrischen Netzes. Das bedeutet, dass die vorgestellten Modelle nur auf Basis der Simulation lernen. Je besser die Simulation desto besser das Modell. Wie gut eine Simulation ist, ist abhängig von ihrem Detailgrad und der Qualität der verwendeten Daten. Ein weiterer Punkt ist, dass Neuronale Netze als Teil von DQN in ihrer Beschaffenheit nicht erklärbar sind. Bei dem Betrieb der Steuerung ist die Grundlage der Entscheidungen der KI nicht bekannt. [4,5]

Fazit und Ausblick:

Die vorgestellte Methode bietet viel Potential für die Netze der Zukunft. Die vorgestellten Ergebnisse basieren bisher jedoch nur auf Simulationen. Außerdem kann die Steuerung des RL-Systems nicht vollständig nachvollzogen werden, sodass bei Fehlern keine Fehleranalyse gemacht werden kann. Alles in allem hat RL die theoretische Möglichkeit teuren Netzausbau zu vermeiden und die Netzkomponenten optimal zu steuern. Ob die Methode praktisch funktioniert, die Qualität der Trainingsdaten und der Simulation ausreichend ist, muss noch abschließend in Feldversuchen getestet werden.

Referenzen

- [1] H. Kazmi, D. Didden, N. Wiese and J. Driesen, "Sample efficient reinforcement learning with domain randomization for automated demand response in low-voltage grids," in *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, doi: 10.1109/JESTIE.2021.3117119.
- [2] J. Duan *et al.*, "Deep-Reinforcement-Learning-Based Autonomous Voltage Control for Power Grid Operations," in *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 814-817, Jan. 2020, doi: 10.1109/TPWRS.2019.2941134.
- [3] M. Kelker, L. Quakernack, J. Haubrock (2021). Multi agent deep Q-reinforcement learning for autonomous low voltage grid control. in *IEEE PES Innovative Smart Grid Technologies*
- [4] L. Kaiser, M. Babaeizadeh, P. Milos, B. Osinski, R. H. Campbell, K. Czechowski, D. Erhan, C. Finn, P. Kozakowski, S. Levine, et al., Model-based reinforcement learning for atari, arXiv preprint arXiv:1903.00374.
- [5] R. S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction, 2nd Edition, MIT press, 2018