



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Niklas Panten

**Deep Reinforcement Learning zur
Betrieboptimierung hybrider
industrieller Energienetze**

**Schriftenreihe des PTW
„Innovation Fertigungstechnik“**

Herausgeber
Prof. Dr.-Ing. Eberhard Abele
Prof. Dr.-Ing. Joachim Metternich
Prof. Dr.-Ing. Matthias Weigold

PTW
TU DARMSTADT

DEEP REINFORCEMENT LEARNING
ZUR BETRIEBSOPTIMIERUNG
HYBRIDER INDUSTRIELLER ENERGIEKETZE

vom Fachbereich Maschinenbau
an der Technischen Universität Darmstadt

zur

Erlangung des Grades eines Doktor-Ingenieurs (Dr.-Ing.)

genehmigte

D I S S E R T A T I O N

vorgelegt von

Niklas Panten, M.Sc.

aus Aachen

Berichterstatter: Prof. Dr.-Ing. Eberhard Abele

Mitberichterstatter: Prof. Dr.-Ing. Jutta Hanson

Tag der Einreichung: 24. Juli 2019

Tag der mündlichen Prüfung: 02. Oktober 2019

Darmstadt 2019

D17

Schriftenreihe des PTW: "Innovation Fertigungstechnik"

Niklas Panten

**Deep Reinforcement Learning zur Betriebs-
optimierung hybrider industrieller Energienetze**

D 17 (Diss. TU Darmstadt)

Shaker Verlag
Düren 2019

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Zugl.: Darmstadt, Techn. Univ., Diss., 2019

Copyright Shaker Verlag 2019

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-7036-1

ISSN 1864-2179

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren

Telefon: 02421 / 99 0 11 - 0 • Telefax: 02421 / 99 0 11 - 9

Internet: www.shaker.de • E-Mail: info@shaker.de

VORWORT DES HERAUSGEBERS

Im Rahmen der Energiewende sind produzierende Betriebe mit zunehmend verschärften gesetzlichen Vorgaben, steigenden und erheblich schwankenden Energiepreisen sowie einem wachsenden Umweltbewusstsein von Konsumenten konfrontiert. Die vielfältigen Energieflüsse in der Produktion mit einer Fülle an stochastischen Einflussfaktoren sind dabei jedoch komplex vernetzt, dass die Analyse, Auslegung und Betriebsweise häufig nicht optimal erfolgt. Gleichzeitig suchen Unternehmen im Transformationsprozess der vierten industriellen Revolution nach technischen und datengetriebenen Lösungen, ihre Prozesse effizienter zu gestalten. Die Digitalisierung und insbesondere Verfahren der Künstlichen Intelligenz (KI) stellen dabei Schlüsseltechnologien dar, um mitunter durch eine optimierte Betriebsweise von Produktions- und Versorgungssystemen Energieeinsparungen zu erzielen sowie den Bedarf flexibel an die volatile Verfügbarkeit Erneuerbarer Energien anzupassen.

Der Autor demonstriert in der vorliegenden Dissertation eindrucksvoll, dass neuartige Verfahren der KI wie Deep Reinforcement Learning Verfahren komplexe Zusammenhänge erlernen und in konkreten Anwendungsfällen durch eine optimierte Ansteuerung versorgungstechnischer Anlagen signifikante Reduktionen der Energiekosten und Lastspitzen erzielen können. Ausgehend von einer Analyse des Status Quo wird ein breit anwendbares Verfahren konzipiert, welches auf Lernerfahrungen an digitalen Zwillingen aufbaut. Dafür werden zugrunde gelegte Modelle sowie wesentliche Aspekte der Implementierung geschildert und eine kritische Evaluierung des Verfahrens nach umfangreichen Hyperparametervariationen vorgenommen.

Damit wird insbesondere im datengetriebenen Kontext der Industrie 4.0 ein aussichtsreiches Verfahren vorgestellt, das für das industrielle Energiemanagement sowie anschließende Forschungsarbeiten von sehr hoher Relevanz ist.

Darmstadt, im Oktober 2019

Prof. Dr.-Ing. Eberhard Abele

*If you are working on something exciting that you really care about,
you don't have to be pushed. The vision pulls you.*

Steve Jobs

VORWORT DES AUTORS

In Zeiten der Digitalisierung und Energiewende war die Vision und der Antrieb für die vorliegende Arbeit, die während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am *Institut für Produktionsmanagement, Technologie und Werkzeugmaschinen (PTW)* der *Technischen Universität Darmstadt* entstand, schnell identifiziert. Am Puls der Zeit braucht es Technologien, die unsere Gesellschaft zu einem besseren Umgang mit den begrenzten Ressourcen unserer Erde befähigen. Diese Vision ist Kern der Aktivitäten unserer Forschungsgruppe *ETA / Energietechnologien und Anwendungen in der Produktion*, im Rahmen derer die vorliegende Arbeit entwickelt wurde.

Mein besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr.-Ing. Eberhard Abele, dem Leiter des Instituts, für die hervorragenden Rahmenbedingungen für die Forschung, die Unterstützung, die nahezu unermesslichen schöpferischen Freiheiten und das Vertrauen, das er in mich und meine Arbeit gesetzt hat.

Ich möchte mich zudem herzlich bei Frau Prof. Dr.-Ing. Jutta Hanson für ihr fachliches Interesse an der Arbeit und die bereitwillige Übernahme des Koreferats bedanken.

Die dieser Arbeit zugrunde liegenden Untersuchungen wurden im Rahmen der vom Bundesministerium für Wirtschaft und Energie (BMWi) geförderten Projekte *ETA-Fabrik* und *PHI-Factory* durchgeführt. Ich möchte allen Forschungs- und Industriepartnern für kritische Nachfragen, die zahlreichen fachlichen Diskussionen und die praktischen Hinweise während der vielen Projekttreffen danken, die mitunter zum Erfolg dieser Arbeit beigetragen haben.

Mein besonderer Dank gilt den vielen Kollegen für die stets sehr freundschaftliche, humorvolle und angenehme Zusammenarbeit. Die Fertigstellung dieser Dissertation während meiner Rolle als Leiter der Forschungsgruppe wäre niemals geglückt, wenn nicht jede(r) durch die individuelle Leistung zum Erfolg des Teams beigetragen hätte. Hervorheben möchte ich Herrn Mark Helfert, Dr.-Ing. Stefan Seifermann, Dominik Flum und danke für die vertrauensvolle Zusammenarbeit im Führungsteam. Mein Dank gilt aber auch insbesondere Herrn Thomas Kohne und Thomas Weber für die zahlreichen fachlichen Diskussionen und das Feedback zu meiner Arbeit. Herrn Heiko Ranzau möchte ich danken für die wertvollen

Vorarbeiten im Rahmen seiner Abschlussarbeit. Unseren ehemaligen Gruppenleitern Herrn Martin Beck und Dr. Philipp Schraml danke ich für das langjährige Vertrauen und die Unterstützung.

Mein herzlicher Dank gilt außerdem meinen wunderbaren Eltern, die mich in allen Lebenslagen nach Kräften unterstützt und viele wertvolle Erfahrungen erst ermöglicht haben.

Zuletzt gebührt insbesondere dir, Jennifer, mein tiefster Dank! Nicht nur für viele motivierende Worte, sondern auch insbesondere dafür, dass du mir stets verständnisvoll mit viel Geduld den nötigen Freiraum und fortwährend einen Ausgleich zur Arbeit gegeben hast. Deine Unterstützung hat wesentlich zum Gelingen dieser Arbeit beigetragen.

Darmstadt, im Oktober 2019

Niklas Panten

2.3.4	<i>Zwischenfazit – Notwendigkeit einer multikriteriellen Betriebsoptimierung hybrider industrieller Energienetze unter zeitvariablen, stochastischen Randbedingungen</i>	44
2.4	KÜNSTLICHE INTELLIGENZ IN DER PRODUKTION	45
2.4.1	<i>Konzept intelligenter, rationaler Agenten</i>	45
2.4.2	<i>Maschinelles Lernen und lernende Agenten</i>	47
2.4.3	<i>Künstliche Neuronale Netze und Deep Reinforcement Learning</i>	51
2.4.4	<i>Anwendungen in der Produktion</i>	56
3	STAND DER WISSENSCHAFT	61
3.1	METHODIK ZUR SYSTEMATISCHEN LITERATURPRÜFUNG.....	61
3.2	EINGRENZUNG UND KONZEPTUALISIERUNG DES THEMAS.....	62
3.3	LITERATURSUCHE UND -FILTERUNG.....	64
3.4	LITERATURANALYSE UND -SYNTHESE.....	66
3.5	FAZIT ZUM STAND DER FORSCHUNG UND FORSCHUNGSDEFIZIT.....	71
4	ZIELSETZUNG UND VORGEHENSWEISE	75
4.1	FORSCHUNGSZIELE.....	75
4.2	ANFORDERUNGEN UND RAHMENBEDINGUNGEN.....	78
4.3	VORGEHENSWEISE.....	82
5	KONZEPTION	85
5.1	BETRIEBSÜBERGREIFENDES ENERGIECONTROLLINGKONZEPT.....	85
5.1.1	<i>Struktur des unterlagerten Controllers</i>	88
5.1.2	<i>Struktur des Optimierungsmoduls</i>	89
5.2	KI-BASIERTES BETRIEBSOPTIMIERUNGSMODUL DURCH DEEP REINFORCEMENT LEARNING.....	91
5.2.1	<i>Basis der Entscheidungsfindungsstrategie</i>	91
5.2.2	<i>Funktionsapproximation und Training Neuronaler Netze</i>	95
5.2.3	<i>Auswahl zentraler Eigenschaften des DRL-Algorithmus</i>	99
5.2.4	<i>Auswahl und Beschreibung des Deep Reinforcement Learning Algorithmus</i>	108
5.2.5	<i>Netzarchitektur und Datenvorverarbeitung</i>	113
5.2.6	<i>Policy- und Reward-Shaping durch Domänenwissen</i>	114
5.3	ZUSAMMENFASSUNG DES GESAMTKONZEPTES.....	117
6	MODELLIERUNG	119
6.1	ÜBERSICHT DES GESAMTSYSTEMS.....	120

6.2	MODELLIERUNGSGRUNDSÄTZE.....	122
6.3	ALLGEMEINES UND NETZKOMPONENTEN	124
6.3.1	<i>Medien und Materialien</i>	124
6.3.2	<i>Netzknoten</i>	125
6.3.3	<i>Steuerung und PI-Regler</i>	127
6.3.4	<i>Pumpen</i>	128
6.3.5	<i>3-Wege-Ventile</i>	130
6.3.6	<i>Rohrleitungen</i>	132
6.3.7	<i>Wärmeübertrager</i>	133
6.3.8	<i>Thermische Pufferspeicher</i>	134
6.4	VERSORGUNGSTECHNISCHE WANDLUNGSANLAGEN	136
6.4.1	<i>Tauchsieder</i>	136
6.4.2	<i>Brennwertkessel</i>	137
6.4.3	<i>Blockheizkraftwerk</i>	138
6.4.4	<i>Kältemaschine und Wärmepumpe</i>	141
6.4.5	<i>Kühlturm</i>	144
6.4.6	<i>Photovoltaik-Anlage</i>	148
6.4.7	<i>Batteriespeicher</i>	150
6.5	GEBÄUDE UND HEIZUNGS-, LÜFTUNGS-, KLIMATECHNIK (HLK)	153
6.5.1	<i>Gebäude</i>	154
6.5.2	<i>Betonkerntemperierung</i>	158
6.5.3	<i>Lüftungsanlage</i>	159
6.6	STÖRGRÖßEN - FERTIGUNG UND UMWELT	161
6.6.1	<i>Fertigung</i>	161
6.6.2	<i>Umwelt</i>	163
6.7	ENERGIEBEZOGENE BETRIEBSKOSTEN UND REWARDS	163
7	IMPLEMENTIERUNG	169
7.1	AGENTENPROGRAMME (REGLER UND KI-BASIERTER OPTIMIERER).....	169
7.1.1	<i>Auswahl der Entwicklungsumgebung</i>	170
7.1.2	<i>Umsetzung des KI-basierten Optimierungsmoduls</i>	173
7.1.3	<i>Umsetzung des unterlagerten Anlagen-Controllers</i>	175
7.1.4	<i>Schnittstelle zwischen unterlagertem Anlagen-Controller und KI-basiertem Optimierungsmodul</i>	177
7.2	VIRTUELLE AGENTENUMGEBUNG (SIMULATION)	178

7.2.1	<i>Anforderungen an die Simulationssoftware</i>	178
7.2.2	<i>Auswahl der Simulationssoftware</i>	180
7.2.3	<i>Umsetzung einer Simulationsbibliothek</i>	182
7.2.4	<i>Umsetzung des Gesamtmodells</i>	184
7.2.5	<i>FMI-Schnittstelle zum Agentenprogramm</i>	185
7.3	REALE AGENTENUMGEBUNG.....	187
7.3.1	<i>Anforderungen an den technischen Aufbau</i>	187
7.3.2	<i>OPC-UA-Schnittstelle zum Agentenprogramm</i>	188
7.4	GRAPHISCHE BENUTZEROBERFLÄCHE.....	189
7.5	ZUSAMMENFASSUNG ZU EINEM FRAMEWORK ZUR BETRIEBSOPTIMIERUNG VON ENERGIESYSTEMEN	190
8	EVALUIERUNG	193
8.1	VORGEHEN UND RANDBEDINGUNGEN.....	194
8.1.1	<i>Dimensionierung der Anlagen</i>	195
8.1.2	<i>Datensätze für Szenarien</i>	196
8.1.3	<i>Zustands- und Aktionsräume</i>	198
8.1.4	<i>Initialisierung von Zustandsgrößen</i>	200
8.1.5	<i>Hyperparameter des Algorithmus</i>	201
8.1.6	<i>Parameter für Kosten sowie Reward- und Policy-Shaping</i>	202
8.1.7	<i>Referenzbetriebsstrategie</i>	204
8.1.8	<i>Kennzahlen zur Bewertung der Performance</i>	205
8.1.9	<i>Hardware-Spezifikationen des genutzten Hochleistungsrechners</i>	206
8.2	VERSUCHSERGEBNISSE (PHASE 1 – TUNING/TRAINING)	206
8.2.1	<i>Einfluss der Normalisierung des States und der Rewards</i>	208
8.2.2	<i>Einfluss von Policy- und Reward-Shaping</i>	208
8.2.3	<i>Einfluss der Zeitschrittweite</i>	210
8.2.4	<i>Einfluss der Batch Größe</i>	211
8.2.5	<i>Einfluss von Zustands- und Aktionsraum</i>	213
8.2.6	<i>Einfluss von Netzarchitektur und Lernrate</i>	215
8.2.7	<i>Einfluss von Diskontierungs- und GAE-Faktor</i>	218
8.2.8	<i>Feintuning der Lernrate</i>	219
8.2.9	<i>Vergleich zu weiteren Actor-Critic Algorithmen</i>	221
8.2.10	<i>Verhalten bei Modellfehlern und Änderungen der Kostenstruktur</i>	223
8.3	VERSUCHSERGEBNISSE (PHASE 2 - DEPLOYMENT).....	224

8.3.1	<i>Ergebnisse für System A</i>	225
8.3.2	<i>Ergebnisse für System B</i>	230
8.3.3	<i>Ergebnisse für System C</i>	233
8.4	BEURTEILUNG DER WIRTSCHAFTLICHKEIT.....	238
8.5	DISKUSSION UND KRITISCHE WÜRDIGUNG.....	241
9	ZUSAMMENFASSUNG UND AUSBLICK	249
9.1	ZUSAMMENFASSUNG.....	249
9.2	AUSBLICK	251
	LITERATURVERZEICHNIS	257
A	ANHANG	277
	LEBENS LAUF DES AUTORS	289