

Energy Management 4.0 - Reinforcement Learning in der Energieoptimierung

Themenbereiche:	Gebäudeenergiemanagement, Reinforcement Learning, Simulationsgestützte Optimierung
Studierende:	Ramon Stoffel
Betreuungsperson:	Igor Dremelj
Experte:	Thierry Pollet
Auftraggebende:	Green4 AG, Baar
Keywords:	Energieoptimierung, Nachhaltigkeit, Batteriemanagement, Reinforcement Learning, Gebäudemodellierung

1. Aufgabenstellung

Die wachsenden Anforderungen an Energieeffizienz und Nachhaltigkeit in Gebäuden stellen eine grosse Herausforderung dar, insbesondere in Gebirgsregionen mit extremen klimatischen Bedingungen wie den Alpen. Ziel dieser Bachelorarbeit war es, ein System zu entwickeln, das Reinforcement Learning (RL) mit Simulationssoftware kombiniert, um den Energieverbrauch im Restaurant Wildspitz auf 1580 Metern Höhe zu optimieren. Das Restaurant verfügt über diverse Energiesysteme wie Photovoltaik, Batteriespeicher und eine Wärmepumpe. In der Untersuchung wurde evaluiert, ob ein RL-Agent, der auf einem realistischen Gebäudemodell trainiert wird, in der Lage ist, Betriebsstrategien zu erlernen, die den Eigenverbrauch maximieren, Lastspitzen reduzieren und den Netzbezug minimieren. Die grösste Schwierigkeit lag in der ganzheitlichen Integration dieser Technologien unter alpinen Bedingungen, die durch wechselhafte Wetterlagen und hohe Spitzenlasten geprägt sind. Die Arbeit diente als Machbarkeitsstudie für den Einsatz von RL im Gebäudemanagement und als Basis für weiterführende Entwicklungen.

2. Vorgehen / Lösungskonzept

Um eine nahtlose Verbindung von Simulationssoftware und RL zu gewährleisten, folgte das Projekt einem systematischen Ansatz. Im ersten Schritt wurde ein detailliertes Gebäudemodell in SketchUp erstellt, das mithilfe von OpenStudio konfiguriert und in EnergyPlus zur energetischen Simulation eingesetzt wurde. Dadurch konnten die Energiesysteme des Restaurants – darunter Photovoltaik, Solarthermie, Batteriespeicher, Wärmepumpe und Kühlzellen – realitätsnah abgebildet werden. Reale Wetterdaten (EPW) wurden integriert, um standortspezifische Bedingungen abzudecken.

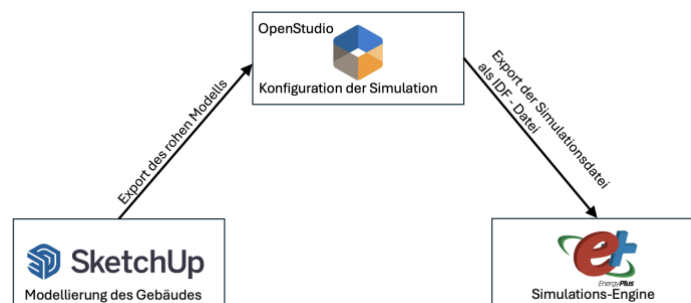


Abbildung 1 - Workflow und Zusammenspiel der integrierten Software

Zeitgleich entstand ein RL-Agent basierend auf dem OpenAI-Gym-Standard und der RLlib-Bibliothek. Er erhielt kontinuierlich Zustandsinformationen wie Batterieladestand, Energieverbrauch und 72-stündige Solarprognosen. Auf dieser Basis entschied der Agent über Lade- und Entladeraten der Batterie und wurde über eine Belohnungsfunktion darauf trainiert, den Netzbezug zu reduzieren und effiziente Betriebsstrategien zu entwickeln. Der genutzte PPO-Algorithmus wurde durch projektspezifische Anpassungen weiter optimiert. Von zentraler Bedeutung war die bidirektionale Kommunikation zwischen EnergyPlus und dem RL-Agenten, die einen Echtzeitaustausch von Zuständen und Aktionen erlaubte. Darüber hinaus wurden dynamische

Trainingsszenarien definiert, um unterschiedliche Startbedingungen und saisonale Einflüsse realistisch abzubilden.

3. Spezielle Herausforderungen

Die Integration von Simulationssoftware und RL brachte verschiedene technische und methodische Herausforderungen mit sich. Zu den grössten Hindernissen gehörten:

- **Kompatibilitätsprobleme zwischen Gym-Versionen:** Frühere und neuere Gym-Implementierungen unterschieden sich in der Handhabung von Episodenendbedingungen, was eine Umstellung des Codes auf aktuelle Konventionen erforderte.
- **Einschränkungen der Rllib-energyplus-Bibliothek:** Die Standardimplementierung erlaubte lediglich die Steuerung eines einzelnen Aktuators. Dies war für die simultane Steuerung von Lade- und Entladeprozessen der Batterie unzureichend. Eine tiefgreifende Modifikation einiger Methoden war notwendig, um mehrere Aktoren gleichzeitig ansteuern zu können.
- **Dynamische Anpassung der Simulation:** Um vielfältigere Szenarien zu ermöglichen, musste die IDF-Datei der Simulation vor jedem Episodenstart dynamisch modifiziert werden, um das Startdatum anzupassen.
- **Komplexität der Simulationssoftware:** Die erstmalige Nutzung von EnergyPlus und OpenStudio erforderte eine umfangreiche Einarbeitung, insbesondere bei der Modellierung spezifischer Energiesysteme wie Batteriespeichern und Kühlzellen.

Diese Herausforderungen wurden durch iterative Anpassungen, Debugging und pragmatische Workarounds gelöst. Die gewonnenen Erkenntnisse flossen kontinuierlich in die Verbesserung des Systems ein.

4. Konkrete Ergebnisse

Das Hauptziel dieser Arbeit – die Entwicklung eines funktionsfähigen Gesamtsystems aus Simulation und RL – wurde erreicht. Der RL-Agent zeigte über 110 Trainings-Episoden hinweg eine deutliche Lernkurve. Während der Agent zu Beginn des Trainings zufällige Entscheidungen traf und geringe Einsparungen von 0.1 % im Vergleich zur Baseline erzielte, konnte er diese Einsparungen bis zur Episode 110 auf 1.1 % steigern. Diese Ergebnisse belegen, dass der Agent in der Lage ist, strategische Entscheidungen zu treffen und sein Verhalten basierend auf den erhaltenen Belohnungen zu optimieren.

Netto – Monatliche & Jahres-Gesamt – episode-00000108-12772

Monat	Baseline (kWh)	RL (kWh)	Diff (kWh)	Abs Diff (%)
2007-01	10956.36	10955.92	-0.45	0
2007-02	9408.68	9403.12	-5.57	0.06
2007-03	5189.11	5145.53	-43.58	0.84
2007-04	1639.66	1583.35	-56.31	3.43
2007-05	-739.24	-798.69	-59.45	8.04
2007-06	-2069.84	-2137.08	-67.24	3.25
2007-07	-2646.38	-2704.68	-58.3	2.2
2007-08	-1901.84	-1956.87	-55.03	2.89
2007-09	190.48	132.84	-57.64	30.26
2007-10	3591.89	3554.1	-37.79	1.05
2007-11	6543.16	6535.61	-7.54	0.12
2007-12	10236.43	10236.37	-0.06	0
2008-01	2.65	2.65	0	0
Gesamt	40401.13	39952.16	-448.97	1.11

Abbildung 2 - Netto Tabelle Episode 108

Ein weiterer Erfolg war die Anpassungsfähigkeit des Systems an unterschiedliche Wetterbedingungen und saisonale Schwankungen. Die entwickelte Umgebung ermöglicht es, die Batterie flexibel zu steuern und zukünftige Solarerträge in die Entscheidungsfindung einzubeziehen. Obwohl keine vollständige jährliche Bewertung durchgeführt wurde, liefern die episodischen Ergebnisse klare Hinweise auf das Potenzial des Ansatzes.

5. Ausblick

Die in dieser Arbeit entwickelten Ansätze und Methoden bilden eine solide Grundlage für weiterführende Forschung und Anwendungen. In Zusammenarbeit mit der Green4 AG wird angestrebt, das entwickelte System im Restaurant Wildspitz einzusetzen und weiter zu optimieren. Ziel ist es, die im Labor entwickelten Strategien in die Praxis zu übertragen und ein massentaugliches Produkt zu schaffen, das in weiteren Gebäuden eingesetzt werden kann.

Zukünftige Arbeiten könnten sich auf folgende Aspekte konzentrieren:

- Verfeinerung der Simulationsmodelle, einschliesslich detaillierterer Nutzungsprofile und energetischer Randbedingungen.
- Integration zusätzlicher Energiesysteme wie variable Stromtarife und Nutzerbedürfnisse.
- Entwicklung einer umfassenden Evaluierungspipeline, um die erlernten Strategien über längere Zeiträume zu testen.
- Übertragung der Strategien auf reale Hardware und Gebäudeenergiesysteme, um den Transfer von der Simulation in die Praxis zu ermöglichen.

Die Arbeit zeigt, dass Reinforcement Learning und Simulationsumgebungen effektive Werkzeuge für das Energiemanagement in anspruchsvollen Umgebungen sind. Mit weiterer Entwicklung könnte dieser Ansatz einen wichtigen Beitrag zur Energiewende und zur nachhaltigen Gebäudeversorgung leisten.