

5.8 Forschungsprojekt „Big Data für EnBop“

5.8.1 Einleitung

5.8.1.1 Hintergrund und Problemstellung

Gebäude und Anlagen werden heute zunehmend komplexer. Die Vernetzung im Zeitalter des Internet-of-Things (IoT) nimmt stetig zu [1] und die Anforderungen an Energieeffizienz und Nutzerkomfort steigen auch weiterhin [2]. Aktuell verursachen Gebäude mit ihren technischen Anlagen ca. 35 % des Endenergieverbrauchs in Deutschland [3]. Bis zum Jahr 2050 soll ein nahezu klimaneutraler Gebäudebestand erreicht werden, indem der nicht erneuerbare Anteil des Primärenergiebedarfs für die Raumkonditionierung gegenüber dem Referenzjahr 2008 um 80 % gesenkt wird [4]. Zum Erreichen dieses Ziels sind neben der Umstrukturierung von Wärme, Kälte und Strom auf erneuerbare Energieträger weitere erhebliche Maßnahmen zur Schaffung von Effizienz notwendig. In diesem Zusammenhang diskutieren unterschiedlichste Forschungsansätze den energetischen Mehrverbrauch durch nicht-optimal betriebene Anlagentechnik. Etabliert hat sich die Beschreibung des Performance Gaps, also der Abweichung zwischen eigentlich erzielbarem Soll und dem tatsächlich auftretenden Ist.

Moderne Nichtwohngebäude verfügen heute oft über eine Gebäudeleittechnik bzw. Managementbedienebene, die Sensor- und Aktordaten kontinuierlich erfasst [5]. Diese Daten spiegeln den Betrieb des Gebäudes wider. Sie geben Aufschluss über die Funktionsweise des Gebäudes als Gesamtes, aber

auch über einzelne Anlagen und Gewerke. Diese Daten fallen heute im Gebäudebetrieb massenhaft an, jedoch mangelt es an Strategien, diese automatisiert zu bewältigen und insbesondere die Datenmenge bzw. Vielfalt der Daten („Big Data“) stellen hohe Anforderungen an eine nachgelagerte Auswertung über größere Betriebszeiträume.

Im Sinne einer umfassenden energetischen und funktionstechnischen Überprüfung der technischen Anlagen, werden größere Zeiträume ausgewertet und viele der vorhandenen Messkanäle miteinander kombiniert. Dabei werden schnell Datenmengen erreicht, für die klassische Auswerteverfahren und konventionelle Infrastrukturen an Ihre Grenzen kommen.

5.8.1.2 Motivation und Aufgabenstellung

Die heutigen Daten einer Leittechnik können Messwerte von Sensoren und Aktoren, z. B. Temperaturen, Volumenströme, Drücke oder Ventilstellungen, in unterschiedlichen zeitlichen Auflösungen (häufig sekundlich, minütlich, viertelstündlich etc.) erfassen und speichern. Fehlermeldungen von Steuergeräten, Aktoren, wie z. B. Pumpen und ähnliches, werden ebenfalls aufgezeichnet. Die bereits heute große Datenmenge wird sich in Zukunft weiter erhöhen und verdichten.

Eine ganzheitliche Betrachtung dieser Daten mit einem hohen Automationsgrad kann aber nur erfolgen, wenn genügend Kontextinformationen zur Verfügung stehen. Allgemeine Attribute wie zum Beispiel die physikalische Gebäudestruktur, Gebäudestandort, Nutzungszweck, aber auch detaillierte

und zum Teil subjektive Merkmale wie zum Beispiel der individuelle Nutzerkomfort oder die technische Auslegung einzelner Anlagen und deren teils heterogene Regelungsstrategien erschweren eine allgemeingültige Aussage über die Performance eines Gebäudes.

Für die Analyse von Betriebs- und Automationsdaten eignen sich Big Data Technologien bzw. Methoden der künstlichen Intelligenz. Im Rahmen dieses Forschungsvorhabens wird daher ein algorithmischer Big Data Baukasten als informationstechnologischer Rahmenwert entwickelt, mit dessen Hilfe sich verschiedene Fragestellungen rund um die Analyse von Betriebsdaten effizient beantworten lassen. Um Prozesse überhaupt handhabbar zu gestalten, wurde dieser Baukasten nach einem „Teile-und-herrsche-Prinzip“ aufgebaut. Für die Umsetzung wurden sowohl reale Betriebsdaten, die über die Gebäudeautomation von zwei großen Nicht-Wohngebäuden gewonnen wurden, als auch mittels dynamischer Gebäudesimulationsmodelle synthetisch erzeugte Betriebsdaten verwendet. Die Anwendung des Big Data Baukastens wurde im Rahmen der Forschungsarbeit in zahlreichen Experimenten aus den Bereichen Datenaufbereitung, Datenklassifikation, Merkmalanalyse und Fehlererkennung in gebäudetechnischen Anlagen demonstriert.

Das Ziel dieses Forschungsvorhabens ist somit die Entwicklung von skalierbaren und automatisierten Lösungen zur Unterstützung der Minimierung von Performance Gaps, die durch diverse Mängel während der Planung, Errichtung und dem Betrieb von Gebäuden entstehen können. Dadurch soll bestmöglicher Nutzerkomfort bei gleichzeitig hoher

Energieeffizienz gewährleistet und zur energetischen Betriebsoptimierung des Gebäudebestandes beigetragen werden.

5.8.2 Praktischer Teil

5.8.2.1 Methodik und konzeptioneller Aufbau des Big Data Baukastens

Im Rahmen des Forschungsprojektes werden zahlreiche Algorithmen evaluiert, an Beispielen zur Anwendung gebracht und auch selbst entwickelt. Hierbei wurde schnell deutlich, dass sich Gebäude- und Anlagenanalysen nicht als einzelner komplexer Algorithmus realisieren lassen, sondern vielmehr eine Vielzahl unterschiedlicher Verfahren zu einem sinnvollen Zusammenspiel orchestriert werden müssen. Der Big Data Baukasten soll als methodisches Rahmenwerk dienen, um die Wiederverwendbarkeit einzelner algorithmischer Bausteine per Konstruktion zu gewährleisten.

Als Lösungsansatz für die zuvor entwickelten Problemstellungen wurde im Rahmen dieses Forschungsprojektes der so genannte Big Data Baukasten entwickelt [6]. Dieser fokussiert auf die Implementierung algorithmischer Fragestellungen zur automatisierten Wissenserkennung und Herleitung gebäude- und anlagenspezifischer Merkmale und wurde als Demonstrator umgesetzt. Die grundlegende Idee dieses Konzeptes ist die strukturierte Entwicklung einzelner algorithmischer Bausteine, die sich zu komplexen Anwendungsszenarien zusammensetzen und orchestrieren lassen. Der Big Data Baukasten gliedert sich zur besseren Übersicht

in mehrere algorithmische Kategorien, die jeweils zugehörige Themengebiete adressieren: Datenanbindung, Vorverarbeitung, Analysen, Bewertung und Reporting.

5.8.2.2 Anwendungsbeispiele des Big Data Baukastens

In diesem Abschnitt werden vier typische Anwendungsbeispiele des Big Data Baukastens im Gebäude- und Anlagenkontext vorgestellt. Hierbei liegt der Fokus auf typischen Problemstellungen des Gebäude- und Anlagenbetriebs und dem allgemeinen Vorgehen zur Lösung ebendieser. Die gewählten Anwendungsbeispiele dienen hierbei dem besseren Verständnis der Problemdomäne und sollen Möglichkeiten für automatisierte Big Data Analysen aufzeigen. Die nachfolgenden vier Anwendungsbeispiele basieren auf ausgewählten Experimenten, die im Rahmen des Forschungsvorhabens durchgeführt wurden, und stellen einen Querschnitt der insgesamt betrachteten Themen und Experimente dar.

Anwendungsbeispiel 1: Datenaufbereitung

Ein grundsätzliches Problem, das oft im Feld auftritt, ist die Tatsache, dass Daten, die aus Zählern stammen, häufig nicht den Charakteristika entsprechen, die ein korrekt funktionierender Zähler hat (z. B. stetige Steigung, technisch nicht plausible Sprünge, Korrektur von Zählerrücksetzungen, Zähleraustausch usw.). Somit ist es notwendig, eine Korrektur der Rohdaten vorzunehmen und die Zählerdaten in eine einheitliche kumulierende Darstellung zu überführen. Im Rahmen des Big Data Baukastens findet dies-

bezüglich eine ganze Klasse von Verfahren Anwendung. Zu den eingesetzten Verfahren zählen beispielsweise auch die zeitliche und wertemäßige Interpolation von Messwerten zur Eliminierung von Messfehlern oder die Umrechnung von ereignisbasierten Daten (Change-of-Value) in äquidistante und damit vergleichbare Zeitreihen. Je nach Anwendungszweck kann außerdem die Eliminierung von Ausreißern oder die Umrechnung oder Normalisierung von Skalen sinnvoll sein.

Anwendungsbeispiel 2: Klassifikation und semantische Anreicherung unbekannter Daten

In diesem Kontext können mehrere Verfahren eingesetzt werden, wie z. B. die Klassifikation mit Entscheidungsbäumen oder durch Deep Learning sowie die Analyse des Anlagenkennzeichnungssystems (AKS). Der AKS bietet die Möglichkeit, semantische Informationen für Datenpunkte durch Analyse von Datenpunktnamen zu ermitteln. Typischerweise beschreibt der AKS einen konkreten Datenpunkt eindeutig (innerhalb eines Namensraums, z. B. einer Gebäudeleittechnik) und ist aus mehreren Teilen zusammengesetzt, die jeweils eine bestimmte Information transportieren.

Das allgemeine Vorgehen ist beispielhaft in Abb. 5.8-1 skizziert. In einem ersten Schritt wird der AKS (Geb23_Hzg1_HKSüd_VL_Temp) dahingehend geprüft, ob der Datenpunktnamen generell der Struktur des erwarteten Kennzeichnungssystems entspricht (z. B. Verwendung des korrekten Trennzeichens, Länge des Schlüssels usw.). Wenn der Algorithmus zum Ergebnis kommt, dass die grundlegende Struktur des Namens nicht korrekt ist, ergibt

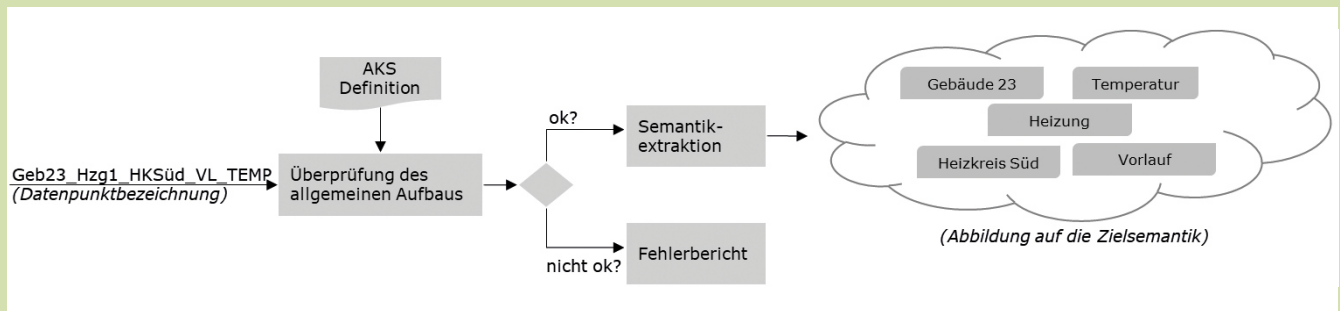


Abb. 5.8-1: Ablauf einer AKS-Analyse

eine weitere Analyse wenig Sinn, weswegen dies an die Aufrufenden rückgemeldet wird. Bei korrektem Aufbau kann mit der Semantikanalyse fortgefahen werden. Diese basiert in der Regel auf dem Abgleich eines vorher festgelegten Schemas, wie das AKS aufgebaut sein muss, und der tatsächlich vorliegenden Datenpunktbezeichnung.

Anwendungsbeispiel 3: Merkmalsidentifikation am Beispiel eines Heizkreises

Das dritte Anwendungsbeispiel soll veranschaulichen, wie mit Hilfe des Big Data Baukastens typische Merkmale technischer Anlagen und Systeme hergeleitet bzw. aus den Rohdaten isoliert werden können. In diesem Zusammenhang wird unter Merkmalen z. B. die Bestimmung der Anlagenaktivität, die Herleitung eines Aktivitätsprofils, die automatisierte Erkennung von Aktivitätsänderungen oder die Herleitung einer Kennlinie für den Heizungsbetrieb verstanden. An dieser Stelle werden exemplarisch die Merkmale des Aktivitätsprofils und der Heizkennlinie erläutert.

Für die Bewertung, ob und wann eine Heizung aktiv war, wurde im Big Data Baukasten ein entsprechender Algorithmus implementiert, der den gesamten Betrachtungszeitraum in so genannte Normwochen (Montag 0 Uhr bis Sonntag 24 Uhr) unterteilt und die jeweiligen Wochenabschnitte aufsummiert. In Abb. 5.8-2 ist dargestellt, wie die zuvor bestimmte Anlagenaktivität in Wochenabschnitte zerteilt

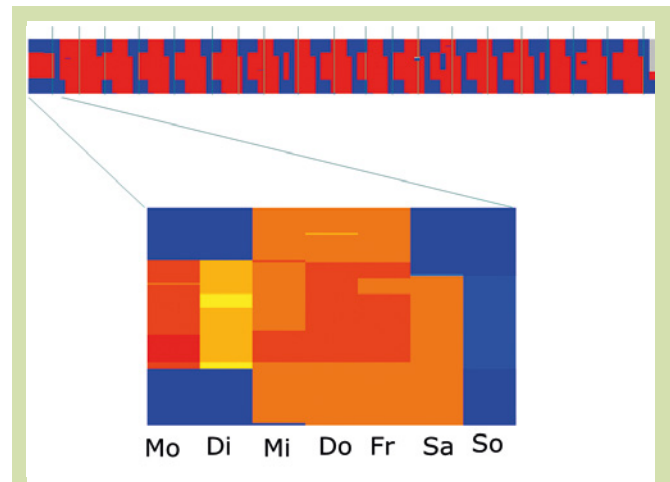


Abb. 5.8-2: Zerlegung in Wochenabschnitte und Aufsummierung zur Normwoche

wird und die jeweiligen Aktivitäten als Normwoche überlagert werden. Je gleichmäßiger die Färbung, desto konsistenter war die Aktivität über die Zeit. Blautöne stehen hierbei für niedrige Aktivität und die Gelbtöne für hohe Aktivität. Die Y-Achse stellt die Uhrzeiten von 0 bis 24 Uhr dar. Am Beispiel ist zu erkennen, dass montags und dienstags früh und spät wenig Aktivität vorlag, jedoch von Mittwoch bis Samstagabend die Anlage weitestgehend durchlief sowie sonntags überwiegend inaktiv war.

Bei der Kennliniendetektion soll diese zunächst anhand der Messdaten des Anlagenbetriebs automatisiert hergeleitet werden, um anschließend eine Validation der Kennlinie, also das Maß der Einhaltung, zu bestimmen. Zunächst werden die Rohdaten als

zweidimensionale Punktwolke, auf der X-Achse die Außentemperatur und auf der Y-Achse die Vorlauf-temperatur im Heizkreis, aufgetragen. Anschließend wird die Punktwolke mit eingeschränktem Wertebereich (Heizgrenztemperatur) und somit einen für den Heizbetrieb repräsentativen Ausschnitt dargestellt. Abb. 5.8-3 zeigt die finale Herleitung einer Regressionsgeraden durch den Algorithmus und die somit ermittelte Heizkennlinie. Im gewählten Beispiel wurde eine Steigung von $-0,4933$ berechnet und die tatsächliche Kennlinie wurde Gebäudeseitig mit der Steigung $-0,5$ betrieben. Auf diese Weise lässt sich mit minimaler Abweichung eine Kennlinie aus den Messwerten herleiten und qualitativ bewerten (hier z. B. Auslegung als Flächenheizung).

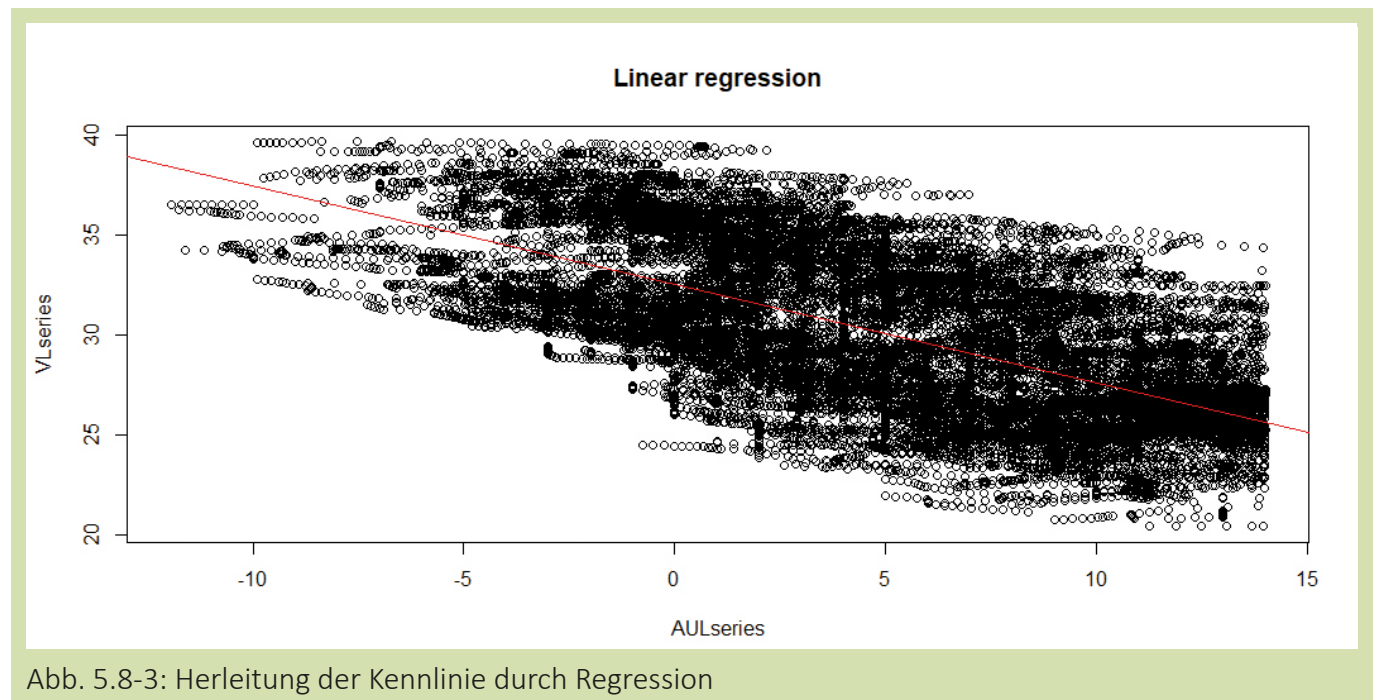


Abb. 5.8-3: Herleitung der Kennlinie durch Regression

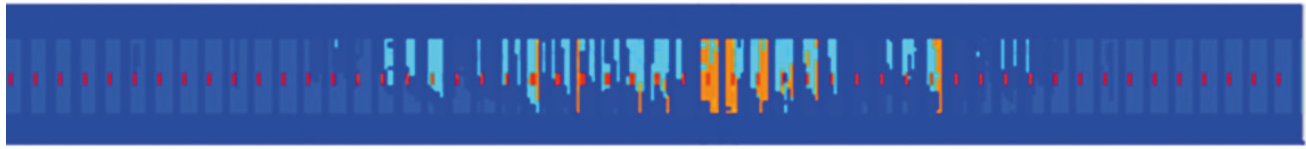


Abb. 5.8-4: Aggregiertes Aktivitätsprofil einer RLT-Anlage



Abb. 5.8-5: Fehlererkennung mit Maschinellem Lernen (Probe mit Realdaten)

Anwendungsbeispiel 4: Automatisierte Fehlererkennung am Beispiel einer RLT-Anlage

Im vierten Anwendungsbeispiel sollen exemplarische Verfahren zur automatisierten Fehlererkennung am Beispiel einer Lüftungsanlage demonstriert werden. In diesem Zusammenhang wird in drei Verfahrensschritten vorgegangen: Dekomposition der Anlage, Definition von Betriebszuständen und Fehlerdefinition, Fehlererkennung.

Raumlufttechnische (RLT) Anlagen bestehen aus unterschiedlichen Sub-Komponenten, zum Beispiel einem Luftvorerhitzer (VE), Luftkühler (KUE), Luftnacherhitzer (NE), Wärmerückgewinnung (WRG) und Ventilatoren (Vent). Für eine möglichst allgemeingültige Betrachtung von RLT-Anlagen ist es sinnvoll, diese zunächst in ihre Sub-Komponenten zu dekomponieren. Dies hat den Zweck, die Komplexität einer Bewertung der Gesamtanlage gemäß dem „Teile-und-herrsche-Prinzip“ in mehrere kleinere Teilprobleme zu zerlegen und diese zunächst

im Einzelnen zu lösen. Durch den Prozess zur Dekomposition sollen abstrakte Anlagen, zunächst nur als Black Box (also ohne Wissen über die innere Struktur) betrachtet werden können.

Im nächsten Verfahrensschritt erfolgt für die Definition von Betriebszuständen und Fehlerdefinition eine zustandsbasierte Analyse dieser Komponenten, welche die Basis für eine algorithmische Fehlererkennung darstellt. Der Systemzustand erfolgt aus der Verkettung einzelner Betriebsmodi auf dieser Komponentenebene. Dabei sind alle möglichen Konstellationen, wie eine RLT-Anlage aufgebaut sein kann, zusammengefasst. Die Gesamtaufstellung der Zustandsmatrix stellt ein sogenanntes „150 %-Modell“ für Vollklimaanlagen dar. Aus den einzelnen Komponentenaktivitäten lässt sich eine Anlagenaktivität herleiten, wie beispielhaft in Abb. 5.8-4 dargestellt und mit unterschiedlichen Farben visualisiert ist.

Für den dritten Verfahrensschritt zur Fehlererken-

nung wurde ein Algorithmus mit Simulationsdaten trainiert, um diesen anschließend auf die Realdaten als Test anzuwenden. Die Übertragung auf Datensätze eines realen Nicht-Wohngebäudes zeigt beispielsweise folgenden Sachverhalt, wie er in Abb. 5.8-5 dargestellt ist. Hier ist zu sehen, dass der identifizierte Betriebsfehler „Gleichzeitiges Heizen und Kühlen“ im gewählten Beispiel sehr häufig auftritt. Um zu validieren, ob es sich tatsächlich um Fehler und nicht um gewünschtes Sollverhalten handelt, wurden im Rahmen der Verfahrensvalidierung die trainierten Daten mit den Realdaten verglichen. Als Fehlerursache konnte ein defektes Kühlerventil für die Probleme und Unregelmäßigkeiten im Gebäudebetrieb identifiziert werden.

5.8.3 Schlussfolgerung

Das Ziel dieses Forschungsvorhabens ist die Entwicklung von skalierbaren und automatisierten Lösungen zur Minimierung von Performance Gaps, die durch diverse Mängel während der Planung, Errichtung und dem Betrieb von Gebäuden entstehen können. Betriebsdaten, die mittels der Gebäudeautomation abrufbar sind, enthalten einen umfangreichen Informationsgehalt, um Probleme in der Betriebsführung von Gebäuden und Anlagen zu identifizieren.

Für die Analyse von Automations- und Betriebsdaten eignen sich Big Data Technologien. Diese zu etablieren stellt jedoch einen erheblichen initialen Aufwand dar. Im Rahmen des Forschungsprojektes wurde daher ein algorithmischer Big Data Baukasten als methodisches informationstechnologisches

Rahmenwerk entwickelt, mit dessen Hilfe sich verschiedene Fragestellungen rund um die Analyse von Betriebsdaten effizient beantworten lassen. Um Prozesse überhaupt handhabbar zu gestalten wurde dieser Baukasten nach dem „Teile-und-herrsche-Prinzip“ sowohl unter Verwendung realer Automationsdaten als auch synthetischer Betriebsdaten mittels Gebäude- und Anlagensimulationen entwickelt. Durch das algorithmische Rahmenwerk lassen sich einerseits individuelle Lösungen umsetzen, die sich durch Modularität, Wiederverwendbarkeit und Kombinierbarkeit auszeichnen. Andererseits ist es auch möglich komplexe Analyseszenarien umzusetzen und zu beherrschen. Die Elemente des Baukastens fokussieren dabei auf Datenkonnektivität, Datenaufbereitung und Vorverarbeitung sowie Datenanalyse und Ergebnisbewertung. Durch die Anwendung des Baukastens wurde eine skalierbare und auf viele Situationen der Qualitätssicherung zugeschnittene Methodik entwickelt. Damit konnten in den Untersuchungsobjekten zahlreiche Betriebsfehler gebäudetechnischer Anlagen identifiziert und Potentiale zur Verbesserung der Gebäudeperformance ausfindig gemacht werden. Zukünftig wird das entwickelte Verfahren zur Minimierung von Performance Gaps und zur energetischen Betriebsoptimierung des Gebäudebestandes beitragen.

Literatur:

- [1] J. Mauerer et. al. Studie Internet of Things 2019. IDG Business Media GmbH, 2018.
- [2] Waide, Paul; Ure, Jim; Smith, Graham; Nordass, Bill. The scope for energy and CO₂, savings in the EU through the use of building automation, 2014
- [3] Bundesministerium für Wirtschaft und Energie. Energieeffizienzstrategie Gebäude: Wege zu einem nahezu klimaneutralen Gebäudebestand. 2015. Webseite <https://www.bmwi.de/Redaktion/DE/Publikationen/Energie/energieeffizienzstrategie-gebaeude.html>.
- [4] Umweltbundesamt. Klimaneutraler Gebäudebestand 2050: Energieeffizienzpotenziale und die Auswirkungen des Klimawandels auf den Gebäudebestand. FKZ 3713 49 101/3716 41 110, Nov. 2017. Webseite <https://www.umweltbundesamt.de/publikationen/klimaneutraler-gebaeudebestand-2050-0>.
- [5] B. Müller, O. Zeidler, M. Geier, A. Badura, W. Müller et. al. Leitfaden zur Gebäudeautomation in öffentlichen Gebäuden zur energetischen Optimierung und Effizienzsteigerung gebäudetechnischer Anlagen. Senatsverwaltung für Stadtentwicklung und Wohnen, 2018.
- [6] J. Mehnert, D. Reiß, S. Plesser, M. Hannen: An algorithmic module toolkit to support quality management for building performance. In: E3S Web Conf. 111, S. 5002, 2019.

Kontakt:

Technische Universität München
Lehrstuhl für Gebäudetechnologie und
klimagerechtes Bauen

Lukas Lauss, M.Sc.
lukas.lauss@tum.de

Dipl.-Ing. Martin Heißler
martin.heissler@tum.de

Prof. Dipl.-Ing. Thomas Auer
thomas.auer@tum.de

Steinbeis-Innovationszentrum energie+
an der Technischen Universität Braunschweig

Jan Mehnert, M.Sc.
j.mehnert@tu-braunschweig.de

Dr. rer. nat. Dirk Reiß

RWTH Aachen
Lehrstuhl für Software Engineering

Sebastian Stüber, M.Sc.
stueber@se-rwth.de

