

Datensets für Demand-Side-Management – Literatur-Review-Basierte Analyse und Forschungsagenda

Dennis Behrens¹, Thorsten Schoormann² und Ralf Knackstedt³

Abstract: Demand-Side-Management (DSM) beschreibt die Bestrebung, das Stromnetz bzw. die Lastverteilungen ausbalanciert zu gestalten. Zur Realisierung von DSM existieren verschiedene Ansätzen, die bislang nur unzureichend evaluiert und verglichen wurden. Als Voraussetzung sowohl für das Evaluieren als auch das Vergleichen dieser Ansätze sind jedoch geeignete Inputdaten, sog. Datensets, nötig. Um die Eignung bestehender Datensets zu prüfen, werden zunächst Anforderungen an DSM-Datensets aus existierenden Ansätzen abgeleitet und konzeptualisiert. Die Klassifizierung der Anforderungen stellt die Basis für das anschließende Literatur Review, bei dem insgesamt 17 relevante Datensets identifiziert wurden. Die Analyse der Datensets zeigt, dass Restriktionen hinsichtlich der Nutzerpräferenzen in keinem Datenset berücksichtigt werden. Im Rahmen der Forschungsagenda werden hierfür Lösungsansätze diskutiert.

Keywords: Demand-Side-Management, Datensets, Literatur-Review

1 Motivation

Das DSM-Konzept zur balancierten Verteilung von Lasten in Stromnetzen existiert bereits seit den achtziger Jahren [Ge85] und wird seither immer populärer. Grund hierfür sind u. a. erneuerbare Energien, die in das existierende Stromnetz integriert werden müssen und Elektroautos, die einen zusätzlichen Energiebedarf erzeugen, aber auch Potentiale bieten. Durch die in Deutschland seit 2010 in allen Neubauten vorgeschriebenen Smart Meter, ist es möglich, einzelne Lasten zu analysieren und zu steuern, sodass ein flexibles Lastmanagement realisiert werden kann bzw. der Nutzer weniger (aktiv) eingreifen muss (s. 21b Abs. 3a EnWG). Durch diese Entwicklungen ist DSM neben dem ursprünglichen industriellen Kontext auch im privaten Bereich anwendbar.

Zur Umsetzung von DSM existiert eine Vielzahl von Ansätzen, die durch z. T. starke Unterschiede geprägt sind [Ko13]. Diese Ansätze werden nicht oder nur unzureichend evaluiert, sodass die Effizienz nicht nachgewiesen werden kann [Ge15]. Zur Evaluation werden meist Datensets genutzt. Um der Realität möglichst zu entsprechen, werden

¹Universität Hildesheim, Informationssysteme und Unternehmensmodellierung, 31141 Hildesheim, dennis.behrens@uni-hildesheim.de

²Universität Hildesheim, Informationssysteme und Unternehmensmodellierung, 31141 Hildesheim, thorsten.schoormann@uni-hildesheim.de

³Universität Hildesheim, Informationssysteme und Unternehmensmodellierung, 31141 Hildesheim, ralf.knackstedt@uni-hildesheim.de

vorzugsweise reale Datensets genutzt [Ba14]. Diese unterscheiden sich jedoch hinsichtlich verschiedener Faktoren (z. B. [Mo14]).

Daher ist es nötig, geeignete Datensets zu nutzen, um DSM-Methoden analysieren, evaluieren und miteinander vergleichen zu können. Hier ergibt sich die Frage, welches Datenset für welche Methode und in welchem Kontext genutzt werden kann? Um diese Frage zu beantworten, wird zunächst der State of the Art der bestehenden Datensets analysiert, da diese zum einen sehr unterschiedlich sind und zum anderen nicht alle Anforderungen der DSM-Methoden erfüllen. Es kann daher die folgende Forschungsfrage abgeleitet werden: **Welche DSM-Datensets sind verfügbar und welche Anforderungen von DSM-Methoden werden von ihnen erfüllt?**

Zur Beantwortung werden in einem ersten Schritt bisherige Arbeiten betrachtet (Abschnitt 2). Anschließend wird das methodische Vorgehen vorgestellt (Abschnitt 3), welches eine Konzeptualisierung von DSM-Methoden vorsieht, bei der vorhandene Anforderungen von DSM-Methoden an Datensets identifiziert werden (Abschnitt 4). Darauf aufbauend wird ein Literatur-Review durchgeführt (Abschnitt 5) und die Resultate analysiert und diskutiert (Abschnitt 6). Abschnitt 7 beinhaltet die daraus abgeleitete Forschungsagenda.

2 Related Work

2.1 Datensets im privaten DSM-Kontext

Eine erste Analyse von existierenden DSM-Datensets wurde von [Mo14] durchgeführt. Sie beschreiben 13 Datensets, von denen ein Datenset selbst aufgezeichnet wurde. Dabei wurden die Lokalität und der Zeitraum, Häuser- und Sensorenanzahl, „Datenset-Features“, wie z. B. Blind- und Scheinleistung, sowie die Auflösung der Aufzeichnung – die Abtastrate – erhoben. Viele dieser Datensets sind nicht mit dem Ziel der Evaluation von DSM-Methoden aufgezeichnet worden, sondern finden sich im Non-intrusiv Load Monitoring (NILM) Kontext wieder. NILM bzw. die NILM-Community befassen sich primär mit der Identifikation von einzelnen Lasten, ohne diese separat aufzuzeichnen [Ha92]. Datensets werden hier genutzt, um Methoden zu evaluieren und zu trainieren.

Eine Übersicht bestehender Datensets, speziell für NILM, findet sich z. B. auch auf dem Blog von [Pa16] wieder, auf dem bekannte Datensets aufgelistet und beschrieben werden. Die Analyse zeigt, dass es keinen Aufzeichnungsstandard gibt. Einen ersten Ansatz – auf Typ- bzw. Metaebene einen Standard zu definieren – liefern [KK14]. Diese beschreiben, mit Hilfe eines Entity-Relationship-Models, ein Metamodell zur Vorgabe benötigter Daten und Strukturen eines Datensets.

2.2 Methoden und Benchmarking im privaten DSM-Kontext

[Ge15] haben einen ersten Überblick erstellt, bei dem 55 DSM-Ansätze, geordnet nach genutzten Methoden und Frameworks, analysiert wurden. Die Autoren folgen dabei der gängigen Einteilung in dezentral, autonom und zentral gesteuerte Frameworks [Ko13]. Dabei wird deutlich, dass keine einheitliche Evaluation oder Vergleichsinstrumente für die einzelnen Ansätze existieren. Ansätze (a) *evaluieren nicht*, (b) *beweisen, dass eine optimale Lösung gefunden wird* oder (c) *evaluieren mit einem Datenset*. Dabei unterscheiden sich die genutzten Datensets zwischen realen und künstlich-erzeugten. Problematisch sind Machbarkeit und Vergleichbarkeit bei allen Ansätzen.

Mit dieser Thematik hat sich z. B. die NILM-Community bereits beschäftigt [Ba14]. Hier wurde ein Benchmarking auf der Basis verschiedener realer Datensets entwickelt, bei dem mehrere NILM-Methoden miteinander verglichen werden können. Auf die Nutzung künstlicher Datensets wird verzichtet und stattdessen auf reale Datensets zurückgegriffen, da diese *die Realität besser abbilden* ([Ba14]). Eine weitere Problematik die adressiert wird, ist die Heterogenität der Datensets. Durch unterschiedliche Eigenschaften und Aufzeichnungsweisen müssen die Datensets zunächst standardisiert werden. Dazu nutzen [Ba14] das Metamodell von [KK14].

3 Methodisches Vorgehen

Zur Identifikation und anschließender Analyse verfügbarer DSM-Datensets anhand relevanter Eigenschaften und Anforderungen wird im Rahmen dieses Beitrages ein mehrstufiges Verfahren verfolgt (s. Abb. 1).

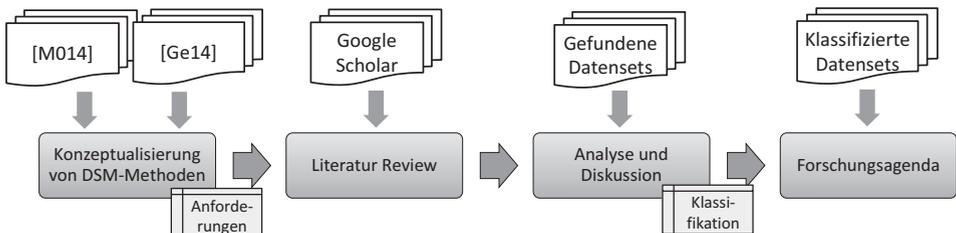


Abb. 1: Übersicht des methodischen Vorgehens

Zunächst werden Anforderungen an DSM-Methoden herausgearbeitet, um darauf aufbauend Dimensionen zur Klassifikation von Datensets abzuleiten. Insbesondere werden dazu die Beiträge von [Mo14] und [Ge15] berücksichtigt, die bereits eine erste Datensetanalyse mit entsprechenden Anforderungen enthalten bzw. eine Übersicht über DSM-Methoden bereitstellen, aus denen weitere Anforderungen identifiziert werden. Das Zwischenfazit der Konzeptualisierung ist die Tabelle „Konzepte“ mit Anforderungsdimensionen und Zuordnung zu DSM-Methoden.

Ziel eines Literatur Reviews ist es, das „große Ganze“ einer Domäne zu verstehen, um irrelevante Forschungsvorhaben zu reduzieren und das bestehende Wissen effektiv nutzen zu können [Br09]. Obwohl das Auswerten der bestehenden Literatur essentiell für die Forschung ist, existieren (a) *teilweise Unklarheiten über die Struktur*, sind (b) *bestehende Vorgehensweisen häufig aus anderen Domänen adaptiert* und (c) *Suchprozesse von gegebenen Guidelines nicht ausführlich beachtet* ([WW02]). Relevant ist vor allem der Suchprozess selbst, weshalb z. B. Literaturfunde, die Suchbegriff-Auswahl und die Evaluation der Ergebnisse präzise dokumentiert werden sollten [Br09]. Aufgrund der methodologischen Rigorosität von Literatur Reviews, folgen wir den fünf Schritten nach [Br09]: (I) *Definition des Forschungsbereichs (Abschnitt 1)*, (II) *Konzeptualisierung von DSM-Methoden (Abschnitt 4)*, (III) *Literatur Review (Abschnitt 5)*, (IV) *Analyse und Diskussion (Abschnitt 6)* sowie (V) *Forschungsagenda (Abschnitt 7)*.

4 Konzeptualisierung von DSM-Methoden

[Mo14] stellen bereits folgende Kriterien von DSM-Methoden vor:

Aufzeichnungsort, Zeitraum und Häuseranzahl. Wo ist das Datenset aufgenommen wurden, wieviele Tage (am Stück) umfasst es und wieviele Wohneinheiten sind enthalten?

Einzelaufzeichnung. Anzahl der einzeln aufgezeichneten Lasten. Entscheidend ist dabei, ob neben dem Gesamtverbrauch die Einzellasten separat aufgezeichnet wurden. [Mo14] geben hier die Anzahl Sensoren und z. T. die Typenanzahl an.

Dimensionierung. Einheit der Lasten. [Mo14] bezeichnen dies als „Features“, da hier u. a. Blind-, Schein- und Wirkleistung getrennt aufgezeichnet werden. Eine eindeutige Angabe in kWh (Wirkleistung) reicht jedoch für die meisten DSM-Methoden aus.

Feingranularität. Integration der Abtastintervalle, d. h. Abstände der Aufzeichnung der Messwerte. Da zwar die Aufwärtskompatibilität jedoch nicht die Abwärtskompatibilität gewährleistet ist, muss eine hohe Abtastrate vorhanden sein. Eine kontinuierliche Optimierung mit DSM ist in der Realität nicht praktikabel und die Taktung kann entsprechend grob sein (z. B. 15 Minuten, wie von vielen DSM-Methoden gefordert).

Bei der Analyse des Datenbedarfs diverser DSM-Methoden, die in [Ge15] aufgeführt sind, konnten die folgenden Anforderungen zusätzlich identifiziert werden:

Erzeugung, Wetterdaten und Speicher. Berücksichtigung der Erzeugung (z. B. PV-Anlagen), um möglichst wenig externen Strom kaufen zu müssen. Kennzeichnung externer Wetterdaten, um z. B. die Auswirkung auf die PV-Anlage(n) zu berechnen oder die Heizleistung abschätzen zu können. Alternativ die Nutzung von Energiespeichern, um z. B. Lastspitzen abzufedern oder Engpässe zu überbrücken.

Flexibilität. Kennzeichnung der zeitlichen Flexibilität von Lasten. Dazu sollten die Lasten getrennt und nicht aggregiert vorliegen. Einige Methoden nehmen unzulässiger Weise pauschal einen Anteil an, der pro Zeiteinheit verschoben werden kann. Eine geeignetere Abbildung wird erzielt, indem die verschiebbaren Lasten getrennt sind und einzeln in das Optimierungsproblem eingehen. Ein Ansatz zur Analyse der Verschiebbarkeit kann bspw. mit einer Referenztabelle erfolgen [KI09].

Trennbarkeit. Kennzeichnung der Trennbarkeit von Lasten. Lasten können zum einen nicht-teilbar sein, d. h. dass die Last während des Auftretens nicht unterbrochen werden kann (z. B. eine Waschmaschine). Andererseits existieren Lasten bei denen dies möglich ist, wie z. B. bei dem Ladevorgang eines Elektroautos. Viele Methoden betrachten daher die Lasten separat und tragen den verschiedenen Lastprofilen Rechnung.

Referenz	Ort	Zeitraum	Anzahl Häuser	Lasten einzeln	Dimensionierung	Feingranularität	Erzeugungs	Flexibilität	Trennbarkeit	Verschiebintervalle	Wetterdaten	Speicher
[Zh13]				X	X	X	X			X	X	
[CC11]				X	X	X	X	X	(X)	X		X
[MM13]		X		X	X	X		X				
[FG12]		X		X	X	X						
[VKM14]		X	X	X	X	X		X		X		
[CYN14]		X	X	X	X	X		X				
[SS14]			X	X	X	X		X		X		
[AYK14]			X	X	X	X		X	X			
[Ma14]				X	X	X		X	X	X		X
[An12b]				X	X	X	X	X				X
[Ji11]				X	X			X		X		
[Du13]			X		X							
[SL14]			X	X	X	X			X	X		
[LF14]				X	X	X	X					
[MHC12]				X	X	X		X		X		
[KG11]				X	X	X		X	X	X		
[MP11]				X	X		X	X	X			
[Al12]				X	X	X		X	X	X		
[Ba14]				X	X	X		X	X			
[Li15]				X	X	X						X
[Mo14b]				X	X							
[HMN12]				X	X	X		X	X	X		
[LPK12]				X	X	X		X				
[Hu14]				X	X	X	X	X				
[Za12]				X	X	X						
[LP13]				X	X	X		X		X		
[BN14]				X	X	X			X			
[Ma12]				X	X	X		X		X		
[AHK13]				X	X	X		X				
[Al13]				X	X	X		X		X		

Tab. 1: Anforderungen der DSM-Methoden

Verschiebungsintervalle. Definition von vordefinierten Anfangs- und Endpunkten zwischen denen eine Last verschoben werden kann. Einige Methoden integrieren daher Nutzerfunktionen. Diese Nutzerfunktionen berücksichtigen zusätzliche Präferenzen der Nutzer, in dem z. B. zusätzliche Kosten bei Missachtung anfallen. Sie enthält Informationen bezüglich des Nutzers und dessen Bedürfnisse, wie z. B. seinen Tagesablauf.

Des Weiteren fokussieren einige Methoden auf Elektroautos oder sog. HVAC-Lasten (Heating, Ventilation and Air-conditioning). Sofern diese keine weiteren Lasten betrachten, wurden die entsprechenden Beiträge nicht betrachtet. In Tab. 1 werden die Resultate dargestellt und den Anforderungen zugeordnet. Die grauen Spalten kennzeichnen neue Anforderungsdimension (vgl. [Mo14]).

5 Literatur Review

Identifikation der Suchbegriffe. Das Ermitteln von Such- und Schlüsselbegriffen im Vokabular der zu analysierenden Domäne ist ein entscheidender Bestandteil des gesamten Suchprozesses [Br09]. Initial wurden daher potenzielle Suchbegriffe verwendet und ausgewertet, um geeignete Items zu identifizieren. In Tab. 2 werden die Begriffe sowie deren Aufkommen in entsprechenden Suchmaschinen dargestellt, graue Zeilen kennzeichnen die ausgewählten Suchbegriffe.

Suchbegriff/-kombination	Google Scholar ⁴	Google	Beschreibung
„Demand-Side-Management“	55.500	466.000	Verknüpfte Items („“).
Demand Side Management	54.200	30.800.000	Auftreten eines Items genügt.
DSM	1.410.000	37.200.000	Abkürz. in anderen Domänen.
DSM „Demand Response“	42.400	423.000	-
NILM	6.960	315.000	-
NIALM	520	12.800	-
„Non-intrusive Load Monitoring“	1.440	20.800	-
“Non-intrusive Appliance Load Monitoring”	529	5.720	-
Dataset „Demand Side Management”	1	32	-
AND “Datensatz”	61	739	-
“s.o.” AND “Datensatz”	2.560	23.200	-
“s.o.” AND “Dataset“	2.840	50.000	-
“s.o.” AND “Data Set”	21.400	240.000	-
Kontext "Demand Side Management" AND "Residential"	16.000	215.000	-
"s.o." AND "Private"	13.800	173.000	-
"s.o." AND "House"			

Tab. 2: Identifikation der Suchbegriffe

⁴ Zugriff auf Google Scholar (<https://scholar.google.de/>) und Google (<http://www.google.de>) am 31.03.2016.

Durch die Evaluation der Begriffe konnten wir folgenden Ausschlusskriterien bestimmen, die nicht zur Beantwortung unserer Forschungsfrage beitragen: „Gas“, „Water“ und „Traffic“ (weitere DSM-Kontexte). Zudem werden Beiträge eliminiert, die zwar die Verwendung von Datensets beschreiben, allerdings nicht das Datenset oder deren Aufzeichnungselemente definieren (z. B. [Ja00]).

Kriterien, die darüber hinaus zum Ausschluss des Datensets geführt haben sind: (a) *Keine aktuelle Verfügbarkeit des Datensets*, (b) *Keine explizite Benennung der aufgenommenen Lastmerkmale* sowie (c) *das Enthalten mindestens eines der Ausschlusskriterien*.

Auswahl der Suchquelle. Da DSM interdisziplinär ist, sind Quellen nötig, die eine breite Analyse ermöglichen. Um eine Vielzahl geeigneter Beiträge zu erhalten, haben wir die Suchmaschine Google Scholar ausgewählt. Durch das Evaluieren verschiedener Suchbegriffskombinationen haben wir folgende Begriffe hergeleitet: *Demand Side Management, Demand Response, NILM, Dataset, Data Set, residential* und *private*.

6 Analyse und Diskussion

6.1 Übersicht und Klassifikation der analysierten Datensets

Datenset	Referenz	Aufzeichnungsart	Zeitraum [in Tagen]	Häuseranzahl	Einzelanzzeichnung	Flexibilität	Trennbarkeit	Versch.-intervalle
ACS-F1	[Gi13]	Schweiz	1	N/A	X	(X)	(X)	
AMPds	[Ma13]	Vancouver, Kanada	360	1	X	(X)	(X)	
BLUED	[An12]	Pittsburgh, USA	8	1	N/A	N/A	N/A	
GREEND	[Mo14]	Österreich, Italien	90	9	X	(X)	(X)	
HES	[Zi12]	Großbritannien	30	251	X	(X)	(X)	
iAWE	[Ba13]	Indien	75	1	X	(X)	(X)	
IHEPCDS	[Li13]	Frankreich	1440	1	(X)	(X)	(X)	
OCTES	[Oc15]	Finnland, Island, Schottland	120	33	N/A	N/A	N/A	
REDD	[KJ11]	Boston, USA	3	6	X	(X)	(X)	
Sample Dataset	[Pe15]	Austin, USA	7	10	X	(X)	(X)	
Smart*	[Ba12] [Um14]	Western Massachusetts, USA	90	1	X	(X)	(X)	
UK-DALE	[Ke15]	Großbritannien	499	4	X	(X)	(X)	
DRED	[Dr15]	Niederlande	61	1	X	(X)	(X)	
ECO	[Ec15]	Schweiz	244	6	X	(X)	(X)	
Energy	[EH16]	Italien	N/A	10	X	(X)	(X)	

@Home							
REFIT	[Mu15]	Glasgow	300 +	20	X	(X)	(X)
TEALD	[Te16]	Burnaby, Kanada	2+	1	X	(X)	(X)

Tab. 3: Klassifikation bestehender DSM-Datensets

Resultierend aus den Anforderungen werden die identifizierten Datensets analysiert, wobei die Anforderungen der Dimensionierung und der Feingranularität als notwendig erachtet und nur solche Datensets berücksichtigt werden, die diese erfüllen. Die Anforderungen nach Wetterdaten, Erzeugung und Speicher, werden nicht berücksichtigt, da diese meist extern aufgezeichnet werden. Heraus sticht das Datenset PLAID [P116], welches zahlreiche reale Lasten bzw. Lastprofile enthält. Diese sind jedoch nicht über einen Zeitraum aufgezeichnet, sondern werden separat und ohne Kontext zur Verfügung gestellt. Es wird daher nicht den „klassischen Datensets“ zugeordnet.

Insgesamt wurden 1.660 Beiträge gefunden, von denen die ersten 250 Einträge berücksichtigt wurden. Durch das Eliminieren von Duplikaten, sowie einer Vor- und Rückwärtssuche konnten im letzten Schritt 17 relevante Beiträge ermittelt werden, die anhand der Anforderungsdimensionen (Abschnitt 4) klassifiziert wurden (Tab. 3). Die Anforderungen werden in den Spalten abgetragen und die Datensets in den Zeilen. Ein „X“ zeigt die explizite, ein (X) die implizite Erfüllung einer Anforderung (s. Abschnitt 6.2). Grau hinterlegte Datensets kennzeichnen neue Datensets [Mo14].

6.2 Diskussion der Ergebnisse

Ort, Zeitraum und Häuseranzahl. Informationen hinsichtlich Ort, Zeitraum und Wohneinheiten sind in allen Datensets enthalten. Ausnahme sind ACS-F1 und Energy@Home, die die Anzahl der Häuser bzw. den Aufzeichnungszeitraum nicht nennen. Aufgrund zahlreicher Unterschiede ist die Vergleichbarkeit der Datensets nicht gegeben.

Einzelaufzeichnung. Fast alle Datensets zeichnen Lasten separat auf. Lediglich OCTES und BLUED aggregieren die Lasten. IHEPCDS berücksichtigt drei verschiedene Lasttypen, sodass nur bedingt von einer ausreichenden Erfüllung gesprochen werden kann.

Trennbarkeit und Flexibilität. Kein Datenset weist explizit auf die Trenn- und Verschiebbarkeit von Lasten hin. Hier kann zwar z. B. über eine Referenztafel [K109] eine Interpretation erfolgen, jedoch wäre eine explizite Kennzeichnung zur Reduzierung des Aufwandes und möglicher Fehler sinnvoll. Bei den Datensets, die die Dimension der Last-Kennzeichnung erfüllen, ist von einer bedingten Erfüllung auszugehen. Bei Lasten, die bereits die vorherige Anforderung nicht erfüllen, kann auch diese Anforderung nicht erfüllt werden, wie z. B. OCTES und BLUED.

Verschiebungsintervall. Keines der Datensets berücksichtigt Verschiebungsintervalle,

wodurch die Angabe von Nutzerpräferenzen wie z. B. in welchem Zeitraum Lasten anfallen dürfen, fehlt. Einige enthalten Zusatzinformationen wie bspw. Bewegungsprofile der Nutzer (z. B. DRED, ECO, iAWE) aus denen sich weitere Daten abgeleitet lassen.

Der Abgleich der bestehenden Datensets mit den Anforderungsdimensionen verdeutlicht, dass Informationsbedarfe bestehen, um Restriktionen aus der realen Welt (besser) berücksichtigen zu können. Vor allem die fehlende Integration von Nutzerpräferenzen in den Datensets stellt eine unzureichend erfüllte Anforderung dar.

6.3 Limitationen

Zunächst wurden im Rahmen dieses Beitrags nur öffentlich-verfügbare Datensets betrachtet, wobei auch nicht-öffentliche Datensets analysiert werden müssen. Diese können jedoch weder für die Evaluation noch für den Vergleich von Methoden genutzt werden, da diese Datensets nicht flächendeckend verfügbar sind. Darüber hinaus ist die geforderte Abbildung der Nutzerpräferenzen sehr aufwendig. Diese Realitätsnähe ist bei einigen Simulationen nicht nötig, da bspw. nur HVAC-Lasten betrachtet werden sollen und somit irrelevant ist, in welchem Intervall Aktionen abgearbeitet werden dürfen. Unabhängig davon sollten die Informationen integriert werden, um entsprechende Funktionen bei der Simulation anzubieten. Die Nutzerakzeptanz, die durch die Nicht-Einhaltung von Nutzerpräferenzen stark sinken würde, ist ein aktuelles Problem des DSM [Pa11].

Das Literatur Review sowie die entsprechenden Suchkonfigurationen basieren auf den Vorgehenshinweisen von [Br09]. Die Auswahl der Suchbegriffe, Suchmaschinen, Ausschlusskriterien sowie die Klassifikation folgen eigenen Entscheidungen, welche Limitationen haben. Es hätten z. B. weitere Schlüsselbegriffe integriert oder weitere Suchmaschinen verwendet werden können, um neue relevante Beiträge zu identifizieren.

7 Forschungsagenda

7.1 Erhebung weiterer Daten

Zur (realitätsnahen) Vervollständigung der Datensets sind neben (I) der Festlegung der Datenbedarfe und Zusatzinformationen, insbesondere (II) die Datenerhebungsmöglichkeiten von großer Bedeutung. Die Erhebungsart muss technische und logistische Details definieren, wie z. B. die Form der Datenaufzeichnung und mögliche Quellen der Daten.

Festlegung weiterer Datenbedarfe (I). Zuerst muss verdeutlicht werden, *welche verschiebbaren Lasten weitere Restriktionen haben.* Hierzu zählen Informationen für die Teil- und Verschiebbarkeit, welche jedoch nachträglich hinzugefügt werden kann.

Schwieriger zu erheben und zu integrieren, ist die *Berücksichtigung der Nutzerpräferenzen*. Es müssen, abhängig von individuellen Präferenzen, die frühesten bzw. spätesten Anfangs- und Endzeitpunkte einer Last dokumentiert werden. Bspw. möchte der Nutzer bei dem Waschvorgang einer Waschmaschine, dass dieser pünktlich zum Ende des Arbeitstages abgeschlossen ist. Daher wäre es sinnvoll, den frühestmöglichen Zeitpunkt als denjenigen zu definieren, an dem die Waschmaschine zu beladen ist. Der spätmöglichste Zeitpunkt wäre entsprechend die geplante Heimkehr abzüglich der Dauer des Waschvorgangs. Allein dieses (einfache) Beispiel zeigt, dass weitere Informationen benötigt werden, wie z. B. Dauer des Waschvorgangs. Eine Möglichkeit wäre, diese Daten sehr detailliert in das Datenset zu integrieren, was jedoch einen Mehraufwand für die Simulation und ein „aufblähen“ des Datensets bedeuten würde. Besser wäre es daher, eine aggregierte Erfassung der Informationen im Datenset, die lediglich die Zeitpunkte enthält.

Festlegung der Datenerhebung (II). Die Form der Aufzeichnung ist weniger schwierig festzulegen und sollte entsprechend den Gegebenheiten angepasst werden. Einige Datensets nutzen hierfür eine Datenbank [Mo14], andere wählen ein simples Dateiformat [Ke15]. Für eine Simulation bietet sich die Bereitstellung der Daten in einfacher Form, z. B. in Textdateien an, da diese vergleichsweise einfach mit Werkzeugen, wie Matlab, eingelesen und verarbeitet werden können. Auch Datenbanken können meist ausgelesen werden, insofern ein entsprechender Auslesevorgang zur Verfügung steht.

Komplexer gestaltet sich die Definition und Auswahl der möglichen Quellen für die Datenerhebung. Hierbei ist insbesondere der Grad der Nutzerintegration entscheidend, um bspw. keine Akzeptanzhemmnisse aufzubauen. Viele Daten sind nötig, um relevante Daten von Anfangs- und Endzeitpunkt generieren zu können. *Ein einfacher und sicherer Weg der Erhebung ist eine direkte Eingabe durch den Nutzer selbst.* Dieser kann via Endgerät, z. B. Smartphone, die benötigten Daten direkt in das System einspeisen. Dies erfordert jedoch einen hohen Aufwand für den Nutzer. Deshalb wäre es von Vorteil, die *Intervalle in denen eine Last verschoben werden kann, durch das System berechnen bzw. prognostizieren zu lassen* und an die individuellen Termine des Nutzers anzupassen. Diese könnten z. B. durch einen Online-Kalender dem System bereitgestellt werden.

Darüber hinaus muss der Nutzer eigene Präferenzen definieren, z. B: Wann trifft der Nutzer zu Hause ein? Wird ein Elektroauto bevorzugt? Neben den Terminen bedarf es zusätzlicher Informationen zur Berechnung. Es wird deutlich, dass ein komplexes Problem und zugleich ein hoher Aufwand zur Integration der Daten bestehen. Da die Simulation dieses nur bedingt leisten kann, entstehen entsprechende Forschungsbedarfe. Um die Simulation mit Daten zu versorgen, bedarf es einer aggregierten Form der Information.

7.2 Generierung von DSM-Datensets

Neben der Möglichkeit, Datensets aufzuzeichnen, können Datensets künstlich erzeugt

werden. Hierbei gibt es verschiedene Vorgehensweisen. Neben der „einfachen“ Möglichkeit, sämtliche Lasten lediglich aneinander zu hängen, wird meist versucht, einem realen Datenset möglichst nahe zu kommen. Dazu sind vor allem die „Verbrauchsprofile“ der einzelnen Lasten wichtig. Zu welchem Zeitpunkt und in welcher Form diese anfallen, bestimmt nicht ein realer Nutzer, sondern der Generator.

Der LoadProfileGenerator bedient sich z. B. einem sog. „desire model“, um das Verhalten von Verbrauchern nachzubilden. Das Modell kann im Generator nahezu beliebig angepasst werden und bietet daher ein weites Nutzer-Spektrum [Pfl16]. Ähnlich gehen auch [Ko14] vor, die mit Hilfe von typischen „load patterns and schedules“ reale Nutzungsprofile nachbilden und zeigen, dass sie (unter bestimmten Bedingungen) ein reales Lastprofil nachbilden können. Auch bei Lastprofilgeneratoren fällt auf, dass wichtige Daten nicht „mitgeloggt“ werden. Selbst wenn es hier ungleich leichter ist dies zu beheben, stellt sich die Frage, woher relevante Daten stammen.

Literaturverzeichnis

- [AHK13] Alam, M. R.; St-Hilaire, M.; Kunz, T.: Cost optimization via rescheduling in smart grids – A linear programming approach. In: Smart Energy Grid Engineering (SEGE), 2013 IEEE International Conference on, 2013.
- [Al12] Ali, S. Q.; Maqbool, S. D.; Imthias Ahamed, T. P.; Malik, N. H.: Pursuit algorithm for optimized load scheduling Power. In: Engineering and Optimization Conference (PEDCO) Melaka, Malaysia, 2012 IEEE International, 2012.
- [Al13] Ali, S. Q.; Maqbool, S. D.; Imthias Ahamed, T. p.; Malik, N. H.: Load scheduling with maximum demand and time of use pricing for microgrids. In: Global Humanitarian Technology Conference: South Asia Satellite (GHTC-SAS), 2013 IEEE, 2013.
- [An12] Anderson K.; Oceau A.; Benitez D.; Carlson, D.; Rowe, A.; Berges, M.: BLUEd: a fully labeled public dataset for Event-Based Non-Intrusive load monitoring research. In: 2nd KDD Workshop on Data Mining Applications in Sustainability, 2012.
- [An12b] De Angelis, F.; Boaro, M.; Fuselli, D.; Squartini, S.; Piazza, F.; Wei, Q.; Wang, D.: Optimal task and energy scheduling in dynamic residential scenarios. In: 9th International Symposium on Neural Networks, China, 2012.
- [AYK14] Ashfaq, A.; Yingyun, S.; Khan, A. Z.: Optimization of economic dispatch problem integrated with stochastic demand side response. In: Intelligent Energy and Power Systems (IEPS), 2014 IEEE International Conference on, 2014.
- [Ba12] Barker S, Mishra A, Irwin D, Cecchet E, Shenoy P, Albrecht J (2012) Smart*: An open data set and tools for enabling research in sustainable homes. In: SustKDD
- [Ba13] Batra, N.; Gulati, M.; Singh, A.; Srivastava, M. B.: It's Different: Insights into Home Energy Consumption in India. In: Proceedings of the 5th ACM Workshop on Embedded Systems for Energy-Efficient Buildings, BuildSys'13. New York. 2013.
- [Ba14] Barta, N.; Kelly, J.; Parson, O.; Dutta, H.; Knottenbelt, W.; Rogers, A.; Singh, A.; Srivastava, M.: NILMTK: An Open Source Toolkit for Non-intrusive Load Monitoring. In: 5th international Conference on Future Energy Systems. Cambridge,

UK. 2014.

- [Ba14] Bassamzadeha, N.; Ghanem, R.; Lu, S.; Kazemtabarm S. J.: Robust scheduling of smart appliances with uncertain electricity prices in a heterogeneous population. In: *Energy and Buildings* Volume 84, December 2014, S. 537–547, 2014.
- [BN14] Bu, H.; Nygard, K. E.: Adaptive Scheduling of Smart Home Appliances Using Fuzzy Goal Programming. In: *ADAPTIVE 2014: The Sixth International Conference on Adaptive and Self-Adaptive Systems and Applications*, 2014.
- [Br09] vom Brocke, J.; Simons, A.; Niehaves, B.; Reimer, K.; Plattfaut, R.; Cleven, A.: Reconstructing the Giant: on the Importance of Rigour in Documenting the Literature Search Process. In: *Proceedings of the 17 European Conference on Information Systems (ECIS)*. Verona, Italy, 161. 2009.
- [CC11] Castillo-Cagigal, M.; Gutierrez, A.; Monasterio-Huelin, F.; Caamano-Martin, E.; Masa, D.; Jimenez-Leube, J.: A semi-distributed electric demand-side management system with PV generation for self-consumption. In: *Energy Conversion and Management* 52 (2011) 2659–2666. 2011.
- [CYN14] Chavali, P.; Yang, P.; Nehorai, A.: A Distributed Algorithm of Appliance Scheduling for Home Energy Management System. In: *Smart Grid, IEEE Transactions on* (Volume: 5, Issue: 1), 2014.
- [Dr15] DRED: Dutch Residential Energy Dataset. <http://www.st.ewi.tudelft.nl/~akshay/dred/>. Stand: 27.09.2015.
- [Du13] Dusparic, I.; Harris, C.; Marinescu, A.; Cahill, V.; Clarke, S.: Multi-agent residential demand response based on load forecasting. In: *Technologies for Sustainability (SusTech), 2013 1st IEEE Conference on*, 2013.
- [Ec15] ECO: Electricity Consumption & Occupancy. <http://www.vs.inf.ethz.ch/res/show.html?what=eco-data>. Stand: 27.09.2015.
- [EH16] Energy@Home: <http://www.energy-home.it/>. Stand: 01.05.2016.
- [FG12] Fathi, M.; Gholami, M.: Localized demand-side management in electric power systems. In: *Smart Grids (ICSG), 2012 2nd Iranian Conference on*, 2012.
- [Ge15] Gerwig, C.; Behrens, D.; Knackstedt, R.; Lessing, H.: Demand Side Management in Residential Contexts – A Literature Review. In: *Informatik, Cottbus*, 2015.
- [Ge85] Gellings, C. W.: The concept of demand-side management for electric utilities. In: *Proceedings of the IEEE*, 73(10):1468–1470. 1985.
- [Gi13] Gisler, C.; Ridi, A.; Zujferey, D.; Khaled, O.; Hennebert, J.: Appliance consumption signature database and recognition test protocols. In: *8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA)*. 2013.
- [Ha92] Hart, G.W.: Nonintrusive appliance load monitoring; In: *Proceedings of the IEEE*, vol. 80, no. 12, S. 1870-1891, Dec 1992.
- [HMN12] Huang, Y.; Mao, S.; Nelms, R. M.: Smooth electric power scheduling in power distribution networks. In: *Globecom Workshops (GC Wkshps), 2012 IEEE*, 2012.
- [Hu14] Hu, T.; Chu, X.; Zhang, W.; Liu, Y.: An optimal day-ahead dispatch strategy for deferrable loads. In: *Power System Technology, 2014 International Conference on*, 2014.

- [Ja00] Jardini, J. A.; Tahan, C. M. V.; Ahn, G. S. U.; Figueiredo, F. M.: Daily Load Profiles for Residential, Commercial and Industrial Low Voltage Consumers. In: IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 15, NO. 1, JANUARY 2000.
- [Ji11] Jia, L.; Yu, Z.; Murphy-Hoye, M. C.; Pratt, A.; Piccioli, E. G.; Tong, L.: Multi-scale stochastic optimization for home energy management. In: Computational Advances in Multi-Sensor Adaptive Processing, 2011 4th IEEE International Workshop on, 2011.
- [Ke15] Kelly J (2015) UK Domestic Appliance-Level Electricity (UK-DALE) dataset. <http://www.doc.ic.ac.uk/~dk3810/data/>. Stand: 27.09.2015
- [KG11] Kim, S-J.; Giannakis, G. B.: Efficient and scalable demand response for the smart power grid. In: Computational Advances in Multi-Sensor Adaptive Processing, 2011 4th IEEE International Workshop on, 2011.
- [KJ11] Kolter, J. Z.; Johnson, M. J.: REDD: A Public Data Set for Energy Disaggregation Research. In: SustKDD Workshop on Data Mining Applications in Sustainability, 2011.
- [KK14] Kelly, J.; Knottenbelt, W.: Metadata for Energy Disaggregation. In: 2nd IEEE International Workshop on Consumer Devices and Systems (CDS 2014). Sweden. 2014.
- [KI09] Klobasa, M.: Dynamische Simulation eines Lastmanagements und Integration von Windenergie in ein Elektrizitätsnetz. In: ISI-Schriftenreihe »Innovationspotenziale«, Fraunhofer IRB Verlag. 2009.
- [Ko13] Kosek, A. M.; Costanzo, G. T.; Binder, H. W.; Gehrke, O.: An Overview of Demand Side Management Control Schemes for Buildings in Smart Grids. In: IEEE International Conference on Smart Eergy Grid Engineering. 2013.
- [Ko14] Kong, W.; Dong, Z. Y.; Chen, G.; Joa, Y.: A Rule Based Domestic Load Profile Generator for Future Smart Grid. In: Australasian Universities Power Engineering Conference, 2014.
- [LF14] Liu, P.; Fu, Y.: Construction of multiband uncertainty set for building energy management with uncertain loads and solar power. In: PES General Meeting | Conference & Exposition. 2014.
- [Li13] Lichman, M.: UCI Machine Learning Repository. <http://archive.ics.uci.edu/ml>. Stand: 27.09.2015, 2013.
- [Li15] Liu, G.-R.; Lin, P.; Fang, Y.; Lin, Y.-B.: Optimal Threshold Policy for In-home Smart Grid with Renewable Generation Integration. In: IEEE Transactions on Parallel & Distributed Systems, , no. 1, S. 1, 2014.
- [LP13] Lee, J.; Park, G.-L.: Power load distribution for wireless sensor and actuator networks in smart grid buildings. In: International Journal of Distributed Sensor Networks, 2013.
- [LPK12] Lee, J.; Park, G.-L.; Kim, H.-J.: Multithreaded Power Consumption Scheduler Based on a Genetic Algorithm. In: International Conference, FGCN 2011, Part I, 2012.
- [Ma12] Maqbool, S. D.; Imthias Ahamed, T. P., Ali , S. Q., Pazheri, F. R., malik, N. H.: Comparison of pursuit and ϵ -Greedy algorithm for load scheduling under real time pricing. In: Power and Energy, 2012 IEEE International Conference on, 2012.
- [Ma13] Makonin, S.; Popowich, F.; Bartram, L.; Gill, B.; Bajic, I. V.: AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research. In: Proceeding of the

IEEE Conference on Electrical Power and Energy, 2013.

- [Ma14] Manasseh, E.; Ohno, S.; Mvuma, A.; Yamamoto, T.: Distributed Demand-side management with load uncertainty. In: Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, 2014.
- [MHC12] Miao, H.; Huang, X.; Chen, G.: A Genetic Evolutionary Task Scheduling Method for Energy Efficiency in Smart Homes. In: Int. Review of Electrical Engineering, 2012.
- [MM13] McNamara, P.; McLoone, S.: Hierarchical Demand Response using Dantzig-Wolfe decomposition. In: Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES. IEEE, S. 1–5, 2013.
- [Mo14] Monacchi, A.; Egarter, D.; Elmenreich, W.; D'Alessandro, S.; Tonello, A. M.: GREEND: An energy consumption dataset of households in Italy and Austria. In: IEEE International Conference on Smart Grid Communications, 2014.
- [Mo14b] Morimoto, N.; Fujita, Y.; Yoshida, M.; Yoshimizu, H.; Takiyama, M.; Akehi, T.; Tanaka, M.: A Power Allocation Management System Using an Algorithm for the Knapsack Problem. In: 38th Annual International Computers, Software and Applications Conference Workshops, 2014.
- [MP11] Missaoui, R.; Ploix, S.: Energy Fluxes optimization for PV integrated Building. In: PowerTech, 2011 IEEE Trondheim, 2011.
- [Mu15] Murray, D.; Coleman, M.; Hauxwell-Baldwin R.; Liao, L.; Kane, T.; Stankovic, L.; Firth, S.; Wilson, C.; Stankovic, V.: REFIT: Smart Homes & Energy Demand Reduction. 2015.
- [Oc15] OCTES: Northern Periphery Programme. <http://www.northernperiphery.eu/en/home/>. Abgerufen am 27.09.2015.
- [Pa11] Paetz, A.-G.; Becker, B.; Fichtner, W.; Schmeck, H.: SHIFTING ELECTRICITY DEMAND WITH SMART HOME TECHNOLOGIES – AN EXPERIMENTAL STUDY ON USER ACCEPTANCE. In: 30th USAEE North American Conference Online Proceedings. 2011.
- [Pa16] Parson, O.: <http://www.oliverparson.co.uk/blog>. Abgerufen om 01.05.2016.
- [Pe15] Pecanstreet (2015) Pecan Street. <http://pecanstreet.org/projects/consortium>. Stand: 27.09.2015.
- [Pf16] Pflugradt, N.: LoadProfileGenerator – LoProGen. <http://www.loadprofilegenerator.de/>. Stand: 01.05.2016.
- [Pl16] PLAID: the Plug Load Appliance Identification Dataset. <http://plaidplug.com/>. Stand: 01.05.2016.
- [SL14] Soliman, H. M.; Leon-Garcia, A.: Game-Theoretic Demand-Side Management With Storage Devices for the Future Smart Grid. In: Smart Grid, IEEE Transactions on (Volume: 5 , Issue: 3), 2014.
- [SS14] Song, L.; van der Schaar, M.: Non-stationary demand side management method for smart grids. In: Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on, 2014.
- [Te16] Makonin, S.: TEALD - The Tautological Energy AnaLog Dataset. <http://teald.org/sample/TEALD.tar.gz>. Stand: 01.05.2016.

- [Um14] UMASS: UMassTraceRepository. <http://traces.cs.umass.edu/index.php/Smart/Smart>. Stand: 27.09.2015.
- [VKM14] Verschae, R.; Kato, T.; Matsuyama, T.: A distributed coordination framework for on-line scheduling and power demand balancing of households communities. In: Control Conference (ECC), 2014 European, 2014.
- [WW02] Webster, J.; Watson, R. T.: Analyzing the past to prepare for the future: Writing a literature review. *Management Information Systems (MIS) Quarterly* 26 (2), 3. 2002.
- [Za12] Zang, C.; Qin, X.; Li, X.; Jin, X.; Che, W.: Simulated annealing based fuzzy Markov game energy management for smart grids. In: Control and Decision Conference, 2013 25th Chinese, 2012.
- [Zh13] Zhao, W.; Cooper, P.; Perez, P.; Ding, L.: Cost-Driven Residential Energy Management for Adaption of Smart Grid and Local Power Generation. 2013.
- [Zi12] Zimmermann, J.-P.; Evans, M.; Griggs, J.; King, N.; Harding, L.; Roberts, P.; Evans, C.: Household Electricity Survey A study of domestic electrical product usage. In: Household Electricity Survey and Household energy, 2012.