

Wissensbasierte Tagesrhythmenerfassung und -auswertung in ubiquitären Umgebungen

Martin Franke

28. Juli 2015

Ein an geregelte Tagesabläufe angepasstes Leben erhöht nicht nur das allgemeine Wohlbefinden, sondern wirkt proaktiv auf Gesundheit und Stresslevel. Feste Rhythmen in Aktivitäten im Tagesverlauf bieten Menschen Sicherheit, Struktur und Orientierung im Alltag. Eine wiederholte oder permanente Missachtung dieser Rhythmen kann zu Schlafproblemen bis hin zu chronischer Depression führen. Personen, die ihren festen Rhythmen nachkommen, sind hingegen weniger von diesen Krankheiten betroffen. Um diese Folgen präventiv zu vermeiden oder zu lösen, kann der eigene Tagesablauf manuell erfasst und ausgewertet werden. Dieser aufwendige Prozess erfolgt bislang nur mit wenig, automatisierter Assistenz und kann durch computergestützte Verfahren erleichtert werden. Dabei ist sowohl ein lückenloser Datensatz an Aktivitäten, als auch das Vorhandensein von möglichst heterogener Sensorik von Bedeutung. Durch eine entsprechende automatisierte Erkennung von Tagesrhythmen kann der Tagesablauf assistiert komplettiert werden und die heterogene Sensorik macht Abweichungen in den Routinen des Nutzers deutlich. Diese Informationen werden dann zur aktiven Lebensunterstützung genutzt, indem Metriken aus diesen abgeleitet oder Anomalien erkannt werden.

Einen möglichen Lösungsansatz zur automatisierten Assistenz bieten dafür sogenannte *Human Activity Recognizer Algorithms* (HARA) aus dem Bereich *Ambient Assisted Living* (AAL). Das Ziel dieser HARA ist es, auf Basis sensorischer

Werte und individuellen Vergleichsmustern die aktuelle Aktivität eines Nutzers zu erkennen. Da diese Systeme häufig im Bereich pflegebedürftiger Personen eingesetzt werden, ist die Auswahl der erkannten Aktivitäten jedoch sehr eingeschränkt und bezieht sich im Regelfall auf nicht-erweiterbare, innerhäusliche Aktivitäten aus der Pflege. Darüber hinaus erfolgt bei diesen Systemen eine Auswertung unter der Voraussetzung, dass sich einzelne Tage in ihrem Ablauf nur wenig unterscheiden, was in anderen Anwendungsdomänen zu Problemen führt. Um zu einer weitergehenden Erfassung des Tagesablaufs zu kommen, muss neben der Auswertung sensorischer Werte auch weiteres Wissen einbezogen werden.

Eine weitere Quelle dieses Wissen bietet die Chronobiologie. Dieser Ansatz zeigt, dass Tagesabläufe nicht nur einer Selbstbestimmung unterliegen, sondern durch endo- und exogene Prozesse verändert werden können. Jahreszeitliche Veränderungen, wie Licht- oder Temperaturschwankungen, haben dabei ebenso Einfluss auf diese Rhythmen, wie tageszeitliche Veränderungen im Hormonhaushalt.

Dieses Wissen kann – nach erfolgreicher Formalisierung – für eine wissensbasierte Erfassung des Tagesrhythmus eingesetzt und somit eine passende Auswahl von Sensorik und Aktivitäten gefunden werden.

Außerdem kann empirisch erfasstes Wissen (also typische Rhythmen unterschiedlicher Nutzergruppen) ebenfalls als weitere Quelle für eine Auswertung einbezogen werden. Dazu muss eine Rückgewinnung generierten Nutzerwissens in Kombination mit einem kollaborativen Mechanismus zum Austausch dieses zwischen Nutzern erfolgen.

Diese Wissensquellen maschinenauswertbar zu formalisieren, kombinieren und bestmöglich auszuwerten, stellt eine Herausforderung dieser Arbeit dar, denn heterogene Datenquellen, unvollständige oder *informationslose* Daten erschweren dem HARA die Auswertung. Dazu müssen Methoden des maschinellen Lernens, semantischer Modellierung und Analyse untersucht und gegebenenfalls weiterentwickelt werden. So lassen sich zwar schon seit einiger Zeit mit Modellen wie PIMO

(Personal Information Model), TMO (Task Model Ontology), NCAL (Calendar Ontology). einzelne Bestandteile eines Tagesablaufs – im Sinne eines elektronischen Kalenders – semantisch beschreiben, jedoch existiert keine formale Beschreibung für eine Verknüpfung von Sensorik, Aktivitäten und Tagesabläufen, womit potentiell vorhandenes Wissen aktuell verloren geht. Um dieses Wissen zu erhalten und auszuwerten, muss ein einheitliches Modell und eine wissensbasierte Architektur geschaffen werden, welche eine Informationserfassung aus heterogenen Datenquellen erlaubt.

Eine solche Architektur löst bestehende Limitationen jetziger HARA in mehreren Ebenen. Erstens kann auf diese Weise im Sinne eines allumfassendes Systems heterogene und erweiterbare Sensorik für die Auswertung des Klassifikators verwendet werden. Zweitens wird durch die Rückgewinnung empirischen Wissens die Erkennung über die Laufzeit kontinuierlich verbessert und ein lückenloses Datenset von Aktivitäten wird möglich. Und drittens kann letztendlich durch eine wissensbasierte Architektur aus Tagesablaufabweichungen eine potentielle Ursache (bis hin zu einzelnen Sensorwerten) für eine Anomalie abgeleitet werden.

Aus den erfassten Tagesabläufen, erkannten Anomalien und den potentielle Ursachen ist es der Person (unter möglicher Zuhilfenahme eines Experten) im Nachzug möglich, einen aus dem Takt geratenen Rhythmus wieder zu synchronisieren.

Die weitere Motivation folgt daraus, dass es in den folgenden Jahrzehnten vermehrt zu psychischen Problemen innerhalb der Gesellschaft kommen wird. So wird vorausgesagt, dass bereits heute 27% der Personen in der EU innerhalb ihres Lebens mindestens einmal mit depressiven Belastungen zu kämpfen haben. Ebenso stellt uns auch der demographische Wandel mit steigendem Alter der Bevölkerung vor neue Herausforderungen.

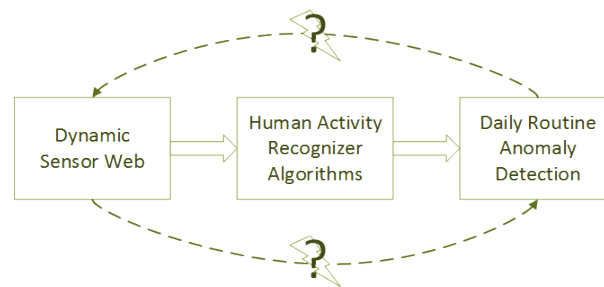
Somit wird es neben der eigentlichen Aktivitätserkennung zunehmend interessant, unter welchen Umständen Abweichungen auftreten, um im besten Fall eine neue *Synchronisierung* des eigenen Rhythmus anzustoßen. Dabei soll das entwickelte

System bestenfalls zwischen einem veränderten Verhalten oder eine Veränderung durch äußere Einflüsse differenzieren und dies dem Nutzer präsentieren. Durch eine frühe Warnung kann die betroffene Person meist selbst ihr Verhalten noch anpassen und zu einem *gesunden* Rhythmus zurückfinden. So ergeben sich Veränderungen, wie z.B. das *Seasonal Affective Disorder* innerhalb eines Jahres. Dies ist vor allem geprägt durch vermehrtes Schlafbedürfnis und einem verringertem Antrieb für sportliche Aktivitäten oder soziale Kontakte. Eine typische Ursache liegt z.B. in einer zu wenig aufgenommenen Menge an natürlichem Tageslicht. Sollte dem System diese Information zugänglich sein, kann diese dem Nutzer zugänglich gemacht werden.

Für ähnliche Anwendungsfälle wurden bereits Systeme entworfen, welche durch Aktivitäts- und Anomalieerkennung dieser automatisierten Assistenz gerecht werden sollen. Dabei geht aktuell Wissen durch die automatisierte Auswertung von Sensoren zu Routinen verloren, da durch die verwendeten Modellierungen oder Verfahren, der Weg nur in eine Richtung durchlaufen werden kann. So existieren Systeme, welche aus Sensornetzen Aktivitäten klassifizieren, welche aber meist nicht in ihren Sensoren nicht zur Laufzeit erweiterbar sind. Außerdem existieren Ansätze, um aus den Aktivitäten Routinen abzuleiten und auf diesen Abweichungen (als potentielle Gefahrensituationen) zu erkennen.

Kein aktuelles System beschäftigt sich jedoch mit der Frage, wie aus dynamischen Sensornetzen, Aktivitäten abgeleitet und auf Basis von Abweichungen Anomalien in den Sensorwerten (also eine potentielle Ursache) gefunden werden kann. Somit ist kein *Tracing* der Daten von Sensorwerten zu den Aktivitäten möglich. Darüber hinaus können aktuelle Systeme nicht nahtlos mit Sensorik erweitert werden, wodurch u.a. Wissen über die Umgebungsfaktoren, wie Außentemperatur und Lichtverhältnisse verloren geht, was ebenfalls laut chronobiologischer Erkenntnisse als wichtige Anforderung gesehen wird, da der Mensch aber stark von diesen Faktoren abhängt.

Um diese Lücke zu schließen, untersucht diese Arbeit die Nutzung des zur Realität werdenden Internet of Things (IoT) in welchem auf immer mehr sogenann-



te *cyber sensors* zurückgegriffen werden kann. Wenn diese Elemente dann auch noch in einer semantischen Beschreibung vorliegen - das sogenannte *Semantic Web of Things* (SWoT) bilden - kann die vereinheitlichte maschinenauswertbare Beschreibung dafür genutzt werden, um heterogene Objekte als *Smart Objects* (d.h. semantisch beschrieben und mit einer Netzwerkschnittstelle) in beliebige Umgebungen zu integrieren. So können dann bereits bestehende Hardware und Services für Datenerfassung rund um das eigene Leben genutzt werden.

Diese Services besitzen als Voraussetzung meist nur die Sensoren des Smartphones oder günstige Consumer-Hardware, was eine Realwelthaftigkeit in assistierten Umgebungen unterstützt. Außerdem können weitere externe Services, wie Erfassung von Klimainformationen (Wetter) nach erfolgreicher Einbindung ebenfalls genutzt werden.

Für diese intelligente, wissensbasierte Vernetzung von heterogenen Informationsquellen in ubiquitären Umgebungen ergeben sich damit die Zielsetzungen:

- (1) Das entwickelte Verfahren unterstützt eine nahtlose, automatisierte Integration (externer) Sensorik. So lässt sich dieses zur Laufzeit mit heterogenen, verteilten Datenquellen erweitern ohne dass Bestandteile des Klassifikators, des Modells oder der Anomaliedetektion angepasst werden müssen. Darüber hinaus ist das Verfahren gegenüber fehlenden oder unvollständigen Sensorwerten robust. Auch die Herausforderung an synchronen Zugriff, Latenz oder unterschiedliche Formate wurde gelöst.

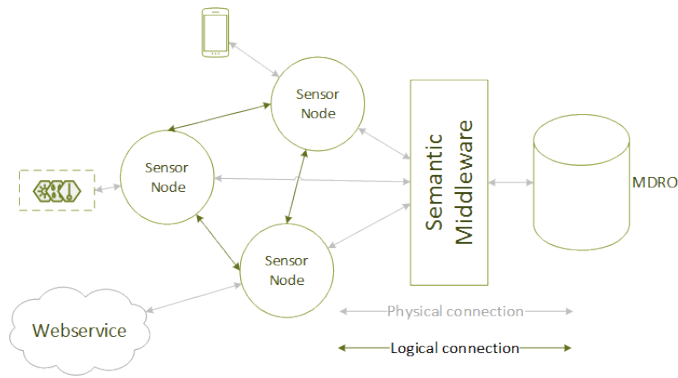
-
- (2) In der praktischen Nutzung spielen die Fragen zu Kosten eine wesentliche Rolle. Es ist wichtig, dass die Anschaffungs- und Wartungskosten in einem akzeptablen Rahmen liegen und somit einer Realwelttauglichkeit entsprechen. Dafür werden nur Sensorik aus dem Bereich Consumer Hardware oder entfernte Dienste (z.B. Wetterdienste) verwendet. Etwaige Ungenauigkeiten dieser Sensorik werden durch die entsprechende Verfahren ausgeglichen. Die Berechnungen können auf *normaler* PC Infrastruktur innerhalb von z.B. maximal 3s für eine Aktivitätsklassifikation durchgeführt werden.
 - (3) Die verwendete Sensorik soll weder störend, noch aufdringlich wirken. Dies war von den jeweiligen Herstellern bereits untersucht wurden und dadurch für die Sensorauswahl gegeben.
 - (4) Alle Daten im System sind semantisch beschrieben. So existiert eine einheitliche Beschreibung für Sensoren, Aktivitäten und Routinen innerhalb der AAL Umgebung.
 - (5) Durch die gemeinsame Verwendung einer einheitlichen semantischen Beschreibung und eines klassifikationbasierten Verfahrens ist es erstmals möglich, angelerntes Wissen adäquat zwischen den Nutzern auszutauschen. Aus diesen Erkenntnissen können gegebenenfalls Gegenmaßnahmen entwickelt, um das Kaltstartproblem eines reinen klassifikationsbasierten Ansatzes einzudämmen.
 - (6) Das entwickelte Verfahren klassifiziert jeden Zeitpunkt eines Tages in einem 15-minütigen Intervall und gibt den Nutzer die Möglichkeit falsch erkannte Aktivitäten effizient zur Laufzeit zu korrigieren. Somit existiert ein lückenloser Tagesablauf, wodurch eine effektive Anomaliedetektion erfolgen kann. Außerdem können nun dadurch dem Nutzer chronobiologische Metriken, wie Work-Life-Balance oder Schlaf-Wach-Rhythmus präsentiert werden. Aus diesen Metriken kann ein aus dem Takt geratenes Leben von einem Experten synchronisiert werden.
 - (7) Die verwendete Ontologie macht es möglich - trotz der Anwendung eines klassifikationsbasierten Verfahrens - aus Aktivitäten und Routinen zu Sensorwerten zurück zu schließen und somit eine Ursache von Anomalien oder Fehlklassifikationen dem Nutzer zu präsentieren. Dadurch entsteht ein Tracing des Wissens von einzelnen Sensorwerten bis zur erfassten Routine aus

Aktivitäten.

- (8) Die verwendete Sensorik soll weder störend, noch aufdringlich wirken. Dies war von den jeweiligen Herstellern bereits untersucht wurden und dadurch für die Sensorauswahl gegeben.
- (9) Der Schutz der Privatsphäre des Nutzers soll nach Möglichkeit weitestgehend gewährleistet werden. Bei dem Verfahren ist es möglich alle Sensorwerte und Berechnungen in der eigenen Infrastruktur getrennt vorzunehmen. Für den Austausch von Nutzerwissen werden durch Pseudonymisierung Daten unkenntlich gemacht, ohne das die Berechnung eingeschränkt ist.
- (10) Die Anpassung des Systems an geänderte Bedingungen findet in allen Bereichen statt. Es kann nahtlos und automatisiert mit Sensorik erweitert werden, welche dann ebenfalls nahtlos für die Aktivitätserkennung genutzt werden. Die Erkennung von Anomalien erfolgt ebenfalls auf diesen erweiterten Sensoren. Somit verbessert sich das System durch Erweiterung zur ohne Anpassung.
- (11) Das Verfahren aus erfassten empirischen Wissen und angelernter Klassifikation erhöht - neben der Wissensmenge (vgl. Metriken Chronobiologie) - die Genauigkeit aktueller HARA signifikant, was in der Evaluation an 12 verschiedenen Personen gezeigt wird.

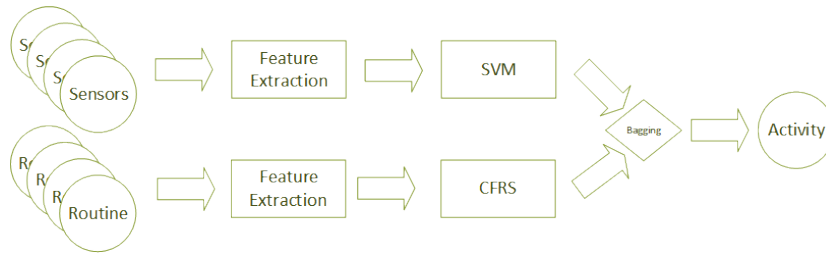
Diese Arbeit stellt somit die erste dar, welche aus dynamischen Sensornetzen (heterogene Sensorik, welche sich zur Laufzeit verändert), hybrider Aktivitätsklassifikation und dem Aufbau von semantischen Routinen, sowohl Abweichungen erkennt, als auch eine signifikante Ursache ermittelt. Dafür wurde aus den Anforderungen der wissensbasierter Systeme die *Modular Daily Routine Ontology* (MDRO) abgeleitet und als zentrale Wissensbasis integriert.

Diese beinhaltet einen Zusammenschluss aus Sensor-, Aktivitäts-, Routine- und Nutzerontologie, welche bereits bestehende Ontologien an benötigten Stellen wiederverwendet. Durch diese *hybride* Struktur von maschinellen Lernen und ontologiebasierten Entscheiden, wurde neben einem zur Laufzeit erweiterbaren Modell auch ein Austausch des Wissens zwischen Nutzern möglich.



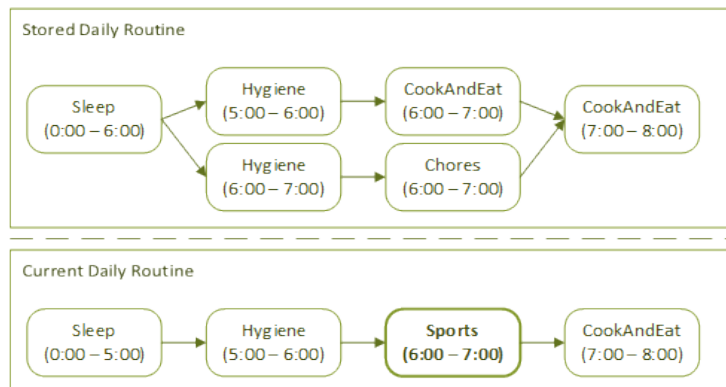
Da diese Arbeit den Fokus auf Rhythmen im Rahmen chronobiologischer Empfehlungen legt, beschränkt sich die Auswahl auf die folgenden Kategorien von Aktivitäten, welche sich sowohl an AAL, als auch an den Erhebungen des statistischen Bundesamtes zu Tagesrhythmen orientiert: *Schlafen; Sport; BildungUndArbeit; EssenUndTrinken; Körperpflege; Hauswirtschaft; Betreuung; Ehrenamt; Freizeit+; Freizeit-*.

Um das dynamische Sensornetz als ubiquitäres System bereitzustellen wurde die *semantische Middleware (SeMiWa)* entwickelt, welche als Koordinator für heterogene Sensoren dient und aus den Anforderungen für *Sensor Web Systeme* abgeleitet ist. Alle Ressourcen (Nodes und Events) werden innerhalb dieser Middleware registriert. Die Middleware wertet eingehende Sensoren und Sensorwerte semantisch auf und verteilt sie entsprechend semantischer Mechanismen. Eine geforderte Dezentralisierung der Middleware wird anhand eines hybriden P2P Verfahrens erreicht, wobei mehrere SeMiWa Server gegenseitig Sensorwerte austauschen können. Alle Producer und Consumer sind, wie in *Event Driven Architecture* vorgesehen, über ein Netzwerkprotokoll voneinander entkoppelt und können asynchron Daten austauschen. Das vereinheitlichte Modell der MDRO und netzwerktransparente Schnittstelle schafft darüber hinaus Interoperabilität und die Möglichkeit heterogene Sensorik gleich welcher Quelle einzubinden. Die Erfassung von Daten ist damit ortsunabhängig und kann sowohl in der häuslichen Umgebung, als auch außerhalb eingesetzt werden. Außerdem ist es möglich Sensoren automatisiert zur Laufzeit zu integrieren, ohne dass das System angepasst werden muss.

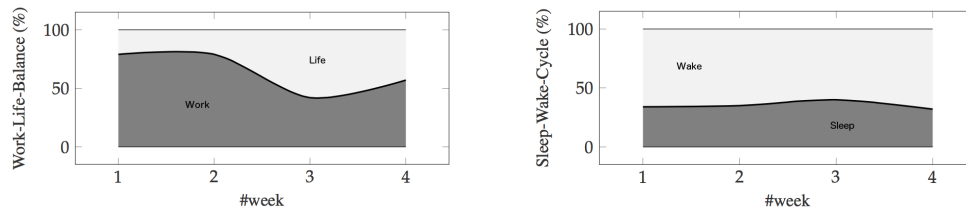


Für die Erkennung der Aktivität des Nutzers wurde ein hybrides Verfahren entwickelt, welches sich an den Anforderungen an HARAs in AAL orientiert. Die Verwendung zwei orthogonaler Verfahren mit Anpassung einer Gewichtung (vgl. *Bagging Algorithms*) nutzt adaptiv das gesamte vorhandene Wissen. Zum einen maschinelles Lernen auf Basis von Sensorwerten, und zum anderen Präferenzen des Nutzers und das kollektive Wissen der Routinen anderer Nutzer. Dies löst damit das Kaltstartproblem durch Vorhandensein von nutzbarer Routinen anderer Nutzer und verbessert sich damit über die Laufzeit.

Auf Basis der erfassten Aktivitäten zu Routinen werden *typische* Rhythmen eines Nutzers abgeleitet und auf einem Graphen zusammengefasst. Für diese Graphen wurde ein Algorithmus entwickelt, welcher Anomalien aufzeigt und die Position (innerhalb des Graphen) dieser Unterschiede ermittelt. Mittels dieser Position kann nun in dem wissensbasierten System eine signifikante Ursache in den Sensorwerten ermittelt werden.



Eine Referenzimplementierung zeigt die Umsetzung der Wissensbasen und Verfahren. Diese wurde zur Evaluierung mit Testdaten und eigens erfassten Daten aus dem



AAL genutzt. Dabei zeigte sich die hohe Genauigkeit des Aktivitätserkennungsverfahrens und die Validität der Ursachenerkennung. Darüber hinaus können die Metriken der Chronobiologie abgeleitet werden.

Als Ausblick wird maßgeblich die Entwicklung einer GUI für das vorgestellte Verfahren angegeben, um eine Studie mit möglichst vielen Personen aus verschiedenen Nutzergruppen zu führen. Die Ideen zu solch einer GUI wurde bereits in eigens veröffentlichten Arbeiten welche z.B. Gamification Elemente zur Motivationssteigerung oder Mixed Realities einbeziehen, gezeigt.

Mittels der Ursachenerkennung kann in weitergehenden Arbeiten außerdem ein Trend angegeben werden, ob sich ein Sensorwert positiv oder negativ entwickelt hat. Diese Idee wurde als *healthyInterval* veröffentlicht, dass z.B. bei Unterschreitung der Schritte oder Schlafdauer die Motivation zu mehr Bewegung erhöht wird.

Zusammenfassend lässt sich sagen, dass mit dem entwickelten Verfahren vielfältige Möglichkeiten durch eine nutzer- und anwendungsfallzentrierte Weiterentwicklung gegeben sind. Durch eine Langzeitbenutzung sollte nun nachgewiesen werden können, dass Folgen unregelmäßiger Rhythmen präventiv vermieden werden.