Im iHomeLab der HSLU wird innovative Gebäudetechnik erforscht.

Der verschwiegene Sprachassistent

Offline-Spracherkennung | Sprachassistenten können das Leben vereinfachen, gerade für Menschen, die sich im Umgang mit Computern nicht sattelfest fühlen. Aber sie werfen auch Fragen zu den Themen Datenschutz und Privatsphäre auf. Das iHomeLab der Hochschule Luzern hat mit der Brelag Schweiz AG den Bodyguard-Sprachassistenten entwickelt, der keine Daten in die Cloud übermittelt.

GUIDO KNIESEL, ANDREW PAICE

Wer hat es nicht schon erlebt: Am Abend hat man sich in Anwesenheit von Siri oder Alexa mit Freunden über Jeans unterhalten, am nächsten Tag erscheint bei der Internet-Suche Werbung über die erwähnte Marke. Zufall? Einbildung? Oder hat sich die Sprachassistentin da doch versehentlich eingeschaltet und Informationen eines privaten Gesprächs weitergeleitet? Diese Unsicherheit begleitet viele Nutzerinnen und Nutzer von Sprachassistenten – nicht ganz unbegründet: Die Mikrofone müssen die Umgebung permanent abhören, damit der Sprachassistent das Aktivierungswort nicht verpasst. «Man kann nie genau wissen, wo die Daten in der Cloud landen und wer Zugriff darauf hat. Auch bietet eine solche Architektur Angriffspunkte für Cyberattacken», gibt Andrew Paice, Leiter des iHomeLab, zu bedenken. Das Kompetenzzentrum der Hochschule Luzern und die Brelag Schweiz AG haben sich deshalb zusammengetan, um einen Sprachassistenten zu entwickeln, der offline funktioniert, also komplett ohne Internetanschluss auskommt. Innosuisse unterstützte das Projekt finanziell. «Bodyguard» heisst das handliche Gerät, das die gesamte Intelligenz in sich vereint, um Absichten und Inhalte aus dem gesprochenen Text zu erkennen und entsprechend zu reagieren.

Ausgangslage des Projekts

Das Smart-Home-Produkt Knockaut der Brelag Schweiz AG zur Gebäudesteuerung und -automation ermöglicht sowohl im privaten Heim als auch in grösseren Bauten und Gewerbeimmobilien die Integration von unterschiedlichen Herstellern, Technologien,

Bild: HSLU

Consumer-Geräten und Diensten in nur einem System bzw. einer App. So lassen sich beispielsweise die Beleuchtung, Jalousien, Kameras, Energie, Musik und vieles mehr individuell steuern.

Um den Wohnkomfort weiter zu steigern, bietet Knockaut durch die Integration des cloudbasierten Alexa-Voice-Assistants die Möglichkeit an, zahlreiche Funktionalitäten im Smart Home mittels Sprache zu steuern. Die Nutzung von digitalen Assistenten über Cloud-Dienste hat allerdings den entscheidenden Nachteil, dass die Daten in die Cloud transferiert werden müssen und zudem auf den Servern der Anbieter gespeichert werden. Deshalb kommt für viele Menschen aufgrund von erheblichen Bedenken bezüglich Sicherheit, Datenschutz und Privatsphäre eine solche Lösung nicht in Betracht, obwohl sie eine Sprachsteuerung gerne nutzen würden.

Sicherheitsrisiken durch cloudbasierte Sprachsteuerung

Ein Hauptproblem der heute verfügbaren Sprachsteuerungen ist die Tatsache, dass sie cloudbasiert sind. Der Markt für Spracherkennung und Smart-Speaker-Geräte wird von Tech-Konzernen wie Amazon und Google dominiert, die Spracherkennung über Cloud-Dienste anbieten. Diese cloudbasierten Ansätze sind jedoch für die Steuerung von Smart-Home-Systemen aufgrund der hohen Anforderungen an Sicherheit, Datenschutz und Privatsphäre nur bedingt brauchbar, da immer wieder neue Sicherheitslücken und Angriffsszenarien identifiziert werden [1,2]. Auch wurden in der Vergangenheit von den Geräten in die Cloud übermittelte Audiodateien systematisch von Mitarbeitern der Hersteller ausgewertet [3].

Die Verwendung von cloudbasierten Smart-Speaker-Lösungen birgt besonders in gewerblichen und industriellen Umgebungen hohe Risiken hinsichtlich Abhörung und Ausspähung von Geschäftsgeheimnissen. Einerseits ermöglichen cloudbasierte Smart-Speaker gezielte (Lausch-)Angriffe von Hackern über das Internet (z. B. für Industriespionage), andererseits kann es durch die versehentliche Aktivierung der Geräte zu ungewollter und ggf. unbemerkter Übermittlung von Daten an unbefugte Personen kommen.

Permanente Abhörung

Die auf dem Markt erhältlichen Smart-Speaker-Geräte wie Amazon Echo oder Google Home befinden sich stets in einem Lauschzustand, da sie für jede Anfrage bzw. jeden Befehl vom Nutzer zuerst über ein Schlüsselwort wie «Alexa» oder «Hey, Google» aktiviert werden müssen. Deshalb hören die Geräte zwangsweise die Umgebung ab und werten alle empfangenen Audiosignale aus, um das Schlüsselwort jederzeit identifizieren zu können. Dies geschieht laut Hersteller zwar lokal auf den Geräten, kann aber zu unbeabsichtigten Aktivierungen der Geräte führen, wodurch es zu ungewollten Datenübermittlungen von privaten und persönlichen Gesprächen und Geräuschen in die Cloud kommen kann [4,5].

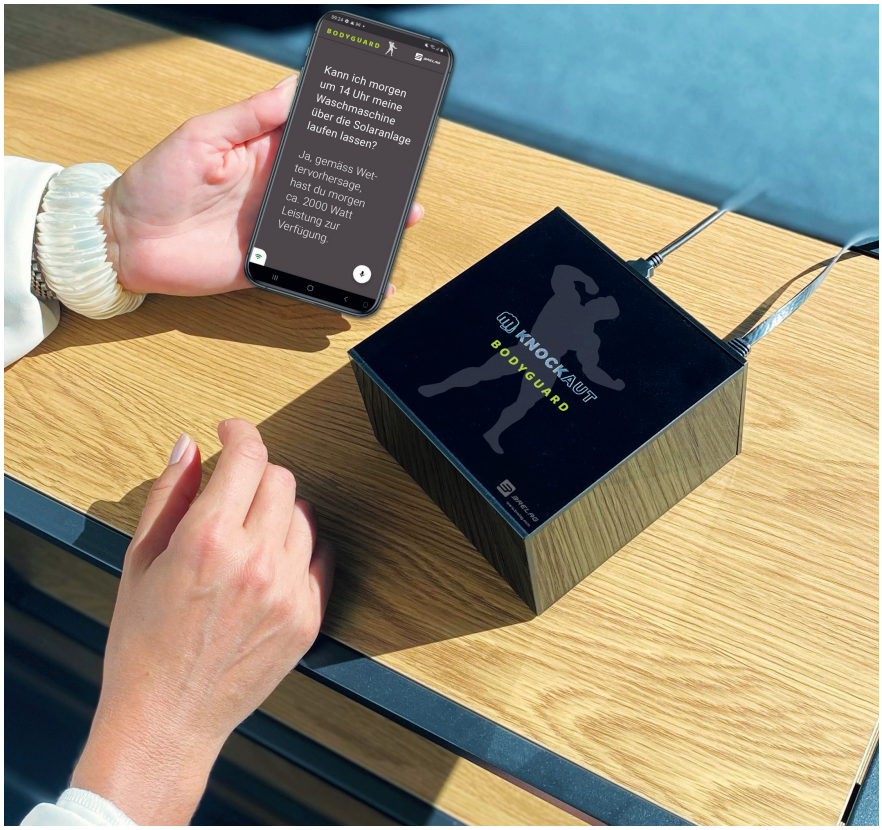
Sprachsteuerungssysteme sind nur reaktiv

Bestehende Sprachsteuerungssysteme in Gebäuden müssen auf Anfragen der Nutzer reagieren, da sie nicht proaktiv von sich aus den Dialog initiieren können. Deshalb müssen sie immer mithören und auf das Aktivierungswort warten.

Eine proaktive Initiierung eines Dialogs, beispielsweise mit anderen Sensorinputs, kann aber nützlich sein, um z. B. bei passenden Gelegenheiten Tipps zu geben oder in bestimmten Situationen Fragen zur Optimierung des Komforts oder der Energieeffizienz zu stellen.

Lösung

Um ein Höchstmass an Privatsphäre zu bieten und so diese Bedenken ausräumen zu können, wurde im Rahmen des Projekts durch das iHomeLab eine lokale Offline-Sprachsteuerung in das System von Brelag integriert. Diese geht nur dann kurz in den Lauschmodus, wenn das System zuvor selbst durch Sensoren aktiviert die Kommunikation zum Bewohner bzw. Gebäudenutzer initiiert hat. Somit hat der Nutzer keine Möglichkeit, von sich aus mit einem Aktivierungswort den Dialog zu initiieren, weil ansonsten das Voice Interface dauerhaft im Lauschzustand sein müsste. Stattdessen darf nur das System selbst in zuvor definierten Situationen (i. d. R. durch Sensoren getriggert) die Kontaktaufnahme per Sprache mit dem Nutzer initiieren, um für die Gebäudesteuerung und das



Die Smart-Home-Gebäudesteuerung Knockaut wurde an der HSLU mit einer lokalen Offline-Sprachsteuerung ausgerüstet und bietet dadurch höchste Privatsphäre.

Bild: Brelag

Energiemanagement relevante Informationen proaktiv zu erfragen. Nur in diesen Situationen – nachdem der Nutzer vom System angesprochen wurde oder der Nutzer selbst manuell über einen Button den Kontakt initiiert – werden die Mikrofone kurz aktiviert und die Umgebung abgehört, um dem Nutzer die Gelegenheit zur Antwort zu geben. Gibt der Nutzer keine Rückmeldung oder wurde der Dialog beendet, werden innerhalb eines bestimmten Zeitfensters (beispielsweise nach 5 s) die Mikrofone wieder deaktiviert und bleiben so lange ausgeschaltet, bis das System selbst erneut eine Kontaktaufnahme per Ansprache initiiert. Mit dieser proaktiven Dialogführung wird das System in die Lage versetzt, durch gezieltes Nachfragen die Steuerung des Smart Home adaptiv auf die Bedürfnisse des Nutzers einzustellen, während gleichzeitig maximale Sicherheit und Privatsphäre gewährleistet bleiben.

Funktionen, Use Cases

«Bodyguard kann zwar prinzipiell auf das Internet zugreifen, um beispielsweise Informationen wie Wetterberichte abzurufen, wird aber ansonsten nur lokal zu klar umrissenen Zwecken eingesetzt», erklärt Paice. Die Einsatzmöglichkeiten sind bereits in der jetzigen Version vielfältig. Zum Beispiel kann Bodyguard über einen CO_2-Sensor feststellen, wenn die Raumluftqualität abnimmt. Dann fragt der Sprachassistent die Nutzerin oder den Nutzer, wie das Smart Home reagieren soll. Er oder sie kann dann zum Beispiel mit einem Sprachbefehl das Fenster öffnen lassen. Brelag bietet Systeme zur Steuerung von Fenstern, Jalousien und Beleuchtung an, die sich mit Bodyguard verbinden lassen.

Bodyguard kann auch eingesetzt werden, um die Produktion einer PV-Anlage vorherzusagen – entweder automatisch oder auf Anfrage. Entsprechend können dann zum Beispiel die Waschmaschine oder der Geschirrspüler auf die Zeiten mit einer hohen Stromproduktion terminiert werden. Dafür braucht das System temporär Zugriff auf das Internet, um Wetterprognosen abzufragen. Die Sprach- und Datenverarbeitung erfolgt jedoch immer lokal; es gelangen keine Informationen des Sprachassistenten ins Internet.

Mittels der Vorhersagen der PV-Produktion kann der Bewohner die Nutzung bzw. die Einschaltzeitpunkte seiner Verbraucher vorausschauend planen. Dies ermöglicht eine Optimierung des Eigenverbrauchs und eine Erhöhung der Autarkiefähigkeit ohne Komforteinbussen. Die Bewohner werden so hinsichtlich Gebäudeeffizienz und ökologischem Fussabdruck sensibilisiert.

Sturzmelder

Aber Bodyguard kann auch als Sturzmelder dienen. Eine Smartwatch mit Beschleunigungssensor erkennt über einen Algorithmus, ob jemand gestürzt ist. Dann schaltet sich das Sprachsystem über die Freisprecheinrichtung der Uhr ein und erkundigt sich: Bist du gestürzt? Brauchst du Hilfe? Wie und wen Bodyguard im Ernstfall alarmieren soll, kann individuell festgelegt werden. «Diese Funktion greift ein für ältere Menschen wichtiges Thema auf: Die Gefahr, zu stürzen und sich dabei zu verletzen. Ich kenne persönlich Menschen, die grundsätzlich noch gut allein hätten leben können. Aber nach einem Sturz blieben sie lange unbemerkt liegen», sagt Andrew Paice. Die Gefahr, zu stürzen, kann dazu führen, dass ältere Menschen ins Altersheim umziehen, obschon sie durchaus noch imstande wären, daheim zu bleiben.

Bodyguard versteht Deutsch mit Schweizer Akzent

Die Mikrofone von Bodyguard sind grundsätzlich ausgeschaltet, ausser wenn sie gewollt aktiviert werden – durch schlechte Luft, einen Sturz oder

RÉSUMÉ

Un assistant vocal discret

Reconnaissance vocale hors ligne pour une protection élevée des données

Les assistants vocaux peuvent simplifier la vie, notamment pour les personnes qui ne se sentent pas très à l'aise avec les ordinateurs, en permettant par exemple de contrôler de nombreuses fonctionnalités de la maison intelligente par la voix. Mais ils soulèvent également des questions liées à la protection des données et de la vie privée, car les solutions de haut-parleurs intelligents basées sur le cloud comportent des risques élevés en matière d'écoute et d'espionnage. C'est pourquoi l'iHomeLab de la Haute école de Lucerne a développé, en collaboration avec Brelag Schweiz AG, l'assistant vocal Bodyguard, qui fonctionne localement et ne transmet aucune donnée dans le cloud. La commande vocale hors ligne ne se met en mode écoute que si le système, activé par des capteurs, a lui-même auparavant initié la communication avec l'habitant ou l'utilisateur du bâtiment. Ce n'est que dans ces situations que les microphones sont brièvement activés pour écouter l'environnement afin de donner à l'utilisateur l'occasion de répondre. Celui-ci n'a donc pas la possibilité d'initier le dialogue de sa propre initiative avec un mot d'activation, car l'interface vocale devrait alors être constamment en mode écoute.

Bodyguard peut en principe accéder à Internet, par exemple pour consulter des informations telles que les bulletins météo, mais n'est sinon utilisé que localement à des fins clairement définies. Par exemple, Bodyguard peut détecter une baisse de la qualité de l'air ambiant grâce à un capteur de CO_2. L'assistant vocal demande alors à l'utilisateur comment la maison intelligente doit réagir, par exemple si elle doit ouvrir les fenêtres par une commande vocale. Bodyguard peut également être utilisé pour prédire la production d'une installation photovoltaïque afin de pouvoir optimiser l'autoconsommation en programmant les consommateurs à des moments où la production d'électricité est élevée. Pour ce faire, le système accède temporairement à Internet pour consulter les prévisions météorologiques. Le traitement de la voix et des données se fait cependant toujours localement.

eine manuelle Eingabe. Und noch einen Vorteil hat Bodyguard: Er versteht auch Deutsch, das mit einem Schweizer Akzent gesprochen wird, und vermeidet so Missverständnisse oder sinnlose Dialoge, wie sie sich mit anderen Sprachassistenten ergeben können.

Fazit

Der entwickelte Prototyp wurde in fünf Testumgebungen in einem Feldtest evaluiert und optimiert. Er verfügt über Algorithmen zur Sprachverarbeitung (Speech-to-Text) in deutscher Sprache und schweizerdeutschem Akzent und Algorithmen zur Extraktion von Intentionen und Inhalten aus unstrukturiertem Text. Algorithmen sorgen zudem dafür, dass Stürze erkannt werden können und die Wetterprognose für die Vorhersage der PV-Produktion genutzt werden kann.

Der Prototyp wurde in bestehende Systemkomponenten (Knockaut-Controller) des Industriepartners integriert. Es stehen zwei Smartphone-Apps zur Verfügung: das Android Voice Interface und das WearOS Voice Interface. Mit Befehlen können Sensoren (Luftfeuchtigkeit, Temperatur, CO_2, Windgeschwindigkeit, Beleuchtungsstärke) in unterschiedlichen Räumen abgefragt und Trigger-Events und Dialoge für die Use Cases CO_2-Überschreitung, Sturzerkennung und PV-Produktionsprognose ausgelöst werden.

Im nächsten Schritt muss der Prototyp noch zu Marktreife gebracht werden, um sicherzustellen, dass er die Kundenbedürfnisse sowie alle rechtlichen und qualitativen Anforderungen optimal erfüllt. Das umfasst beispielsweise die Überarbeitung des Designs für eine kosteneffiziente Herstellung und die Entwicklung einer geeigneten Marketingstrategie, um erfolgreich auf dem Markt platziert zu werden. Ein konkreter Termin, wann das Gerät erhältlich sein wird, steht noch nicht fest.

Referenzen

[1] Fabian Bräunlein, Luise Frerichs, «Smart Spies: Alexa and Google Home expose users to vishing and eavesdropping», Security Research Labs, 2019. www.srlabs.de/blog-post/smart-spies

[2] Deepak Kumar et al., «Skill Squatting Attacks on Amazon Alexa», Usenix Security Symposium, 2018. www.usenix.org/conference/usenixsecurity18/presentation/kumar

[3] Matt Day, Giles Turner, Natalia Drozdiak, «Amazon Workers Are Listening to What You Tell Alexa», Bloomberg, 2019. www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio

[4] Katie Canales, «A couple says that Amazon's Alexa recorded a private conversation and randomly sent it to a friend», Business Insider, 2018. www.businessinsider.com/amazon-alexa-records-private-conversation-2018-5

[5] Edward C. Baig, «Alexa, Google: Parents worry voice assistants eavesdropping on kids», Usatoday, 2019. www.usatoday.com/story/tech/talking-tech/2019/03/28/parents-dont-want-smart-speakers-to-secretly-record-kids-survey/3288806002

Autoren

Guido Kniesel ist Senior wissenschaftlicher Mitarbeiter am iHomeLab.
→ HSLU, 6048 Horw
→ guido.kniesel@hslu.ch

Prof. Dr. **Andrew Paice** ist Leiter des iHomeLab.
→ andrew.paice@hslu.ch

Das Bodyguard-Forschungsprojekt wurde von Innosuisse finanziell unterstützt.