electronics

MDPI

Article

A Framework for Smart Home System with Voice Control Using NLP Methods

Yuliy Iliev [1] and Galina Ilieva [2,*]

[1] Teletek Electronics, 1407 Sofia, Bulgaria
[2] Department of Management and Quantitative Methods in Economics, University of Plovdiv Paisii Hilendarski, 4000 Plovdiv, Bulgaria
* Correspondence: galili@uni-plovdiv.bg

Abstract: The proliferation of information technologies and the emergence of ubiquitous computing have quickly transformed electronic devices from isolated islands of data and control into interconnected parts of intelligent systems. These network-based systems have advanced features, including Internet of Things (IoT) sensors and actuators, multiple connectivity options and multimodal user interfaces, and they also enable remote monitoring and management. In order to develop a human machine interface of smart home systems with speech recognition, we propose a new IoT-fog-cloud framework using natural language processing (NLP) methods. The new methodology adds utterance to command transformation to the existing cloud-based speech-to-text and text-to-speech services. This approach is flexible and can be easily adapted for different types of automation systems and consumer electronics as well as to almost every non-tonal language not currently supported by online platforms for intent detection and classification. The proposed framework has been employed in the development of prototypes of voice user interface extension of existing smart security system via new service for speech intent recognition. Tests on the system were carried out and the obtained results show the effectiveness of the new voice communication option. The speech-based interface is reliable; it facilitates customers and improves their experience with smart home devices.

Keywords: smart system; smart home; voice user interface; speech command recognition; natural language processing; natural language understanding; under-resourced languages

check for updates

Citation: Iliev, Y.; Ilieva, G. A Framework for Smart Home System with Voice Control Using NLP Methods. *Electronics* **2023**, *12*, 116. https://doi.org/10.3390/electronics12010116

Academic Editor: Juan M. Corchado

Received: 30 November 2022
Revised: 21 December 2022
Accepted: 23 December 2022
Published: 27 December 2022

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A home automation system (smart home, smart house) is a communication network comprising of home sensors, devices and appliances (lighting, fans, air conditioners, entertainment systems, surveillance cameras, electronic doors, alarm systems, etc.) for access and monitoring of home environment without human intervention [1]. If smart home components (usually Internet of Things—IoT devices) are connected to the Internet, they can send and receive data from the global network and thus be controlled remotely via different communication protocols [2,3].

Typically, smart home systems employ three different computing layers for data collection: processing and storage—edge, fog and cloud, respectively. The middle layer (fog) of smart home architecture lies between physical IoT devices (edge) and traditional data storage (cloud) levels. This intermediate layer extends the cloud infrastructure and brings computations and storage closer to their data source, i.e., the edge. Distinct from centralised cloud computing infrastructure [4,5], the fog layer consists of multiple nodes which can build a decentralised computing ecosystem. When a fog node receives data, it can decide whether to process it locally or send it to the cloud. In addition, the data can be accessed offline as it can be stored locally. This is another key difference between fog and cloud computing, where data processing and storage are only carried out by remote servers [6–8]. Therefore, the fog layer is more effective than the cloud one in solving

local tasks in real time, since it minimises transmission latency, improves response time, decreases bandwidth consumption and reduces some cyber security risks [9,10].

In this study, we combine sensors and actuators, fog and cloud services in a single information space to extend the capabilities of user interface of smart systems. Human-machine interaction is an important factor for the adoption of new products by consumer electronics market worldwide. On the one hand, smart user interface (SUI) increases sales and improves the corporate image of smart system manufacturers and vendors. On the other hand, an innovative human-machine communication is also useful for customers, as they quickly appreciate the capabilities of new technologies that enter their everyday lives. According to International Data Corporation (IDC), a market research company, worldwide smart home devices market grew by 11.7% in 2021, with double-digit growth forecast as we approach 2026 [11]. The results of Juniper Research study of digital technologies market [12] show that there will be almost 13.5 billion home automation systems in active use by 2025 and voice capabilities are an increasingly common way to control digital devices, linking them into the smart home ecosystem.

In the last decade, the spread of IoT and speech recognition technologies has led to the development of a variety of smart devices that can be controlled from a distance by voice. The advantages of Voice User Interface (VUI) compared to the classical interaction modes are numerous:

1. VUI provides an additional control channel offering hands-free and eyes-free interaction (multi-modality);
2. It enables users to perform multiple actions while communicating (multi-tasking);
3. VUI is analogous to everyday conversation (intuitiveness and ease of use);
4. It is also beneficial to people with physical and cognitive disabilities, who have difficulties interacting with electronic devices through a conventional interface (accessible design) [13].

There are several ecosystems for voice-based technologies that dominate the voice control market such as Google DialogFlow (2014) [14,15], Microsoft Language Understanding (LUIS) (2018) [16,17], IBM Watson Natural Language Understanding (2018) [18,19] and Alibaba Cloud Intelligent Speech Interaction (2020) [20]. These platforms offer a plethora of services, such as transcribing speech to text and vice versa, setting up reminders, searching the Internet and responding to simple queries (weather, traffic or route navigation, playing music or TV) in several languages. The above-mentioned voice service platforms, differing from their predecessors (smart voice speakers and traditional voice assistants). are not limited to specific hardware or operating systems. Instead, they are accessible on any device that is connected to the corresponding cloud platform. The voice platforms also have vertical integrations across multiple industries (retail, transportation, entertainment and media, health) but unfortunately, they have some drawbacks:

1. Speech-to-Text (STT), Text-to-Speech (TTS) and especially natural language understanding (NLU) services are only available for a small subset of the world's 7000+ languages and their variants, and often these services are paid (subscription is needed);
2. Service integration into smart home systems often requires computer science competences (especially in the case of on-premises deployment), a significant amount of time and effort, as well as technical support;
3. The list of recognised commands (skills) is usually domain-specific and needs to be expanded.

The main goal of this research is to develop a new conceptual framework for enhanced voice user interface in smart home systems. The proposed framework incorporates real-time remote monitoring and control through various channels and supports voice interface even for under-resourced languages.

The advantages of the new framework for home automation with voice control are as follows:

1. It allows for the implementation of intelligent user interface (IUI) with "understanding" of domain-specific voice commands in natural language. This option is especially important for under-resourced languages as there is no alternative online service for their automatic intent recognition;
2. The proposed approach is a free of charge alternative to existing paid online NLU services for the most widely used languages;
3. The new speech recognition service for voice control can be deployed in any layer of the IoT-Fog-Cloud (IoTFC) architecture and in the case of positioning in edge or fog level, it can minimise bandwidth loads, improve communication security and increase the intelligent system's efficacy.

The structure of the paper is as follows: in the next section, existing approaches for development of voice user interface for smart devices have been analysed and discussed. The third section presents the features of the proposed new framework for voice control of smart house systems using NLP methods. The next section describes the prototypes implemented to verify the new methodology and discusses its advantages. The last section concludes and outlines our future research directions.

2. Related Work

2.1. Literature Review on IoTFC Architecture

The recent research topics on smart systems architectures can be categorised in two main areas: (1) multi-layer frameworks and business processes for decentralised and real-time data processing and (2) practical implementations of customer electronics systems.

Sun et al. [21] have formulated the computation offloading and resource allocation in general IoTFC architecture as an energy and time cost minimisation problem. Then, they proposed a new algorithm to solve this problem, improving the energy consumption and completion time of application requests. In order to cope with big data and heterogeneity challenges in an IoTFC ecosystem, Chegini et al. [22] designed automatic components for fog resiliency. The advantage of the proposed approach is that it makes the processing of IoT tasks independent of the cloud layer.

Kallel et al. [23] modelled and implemented two IoT-aware business processes. The first one monitors behaviour of children with disabilities to guarantee their safety and facilitate their parents' intervention. The second model facilitates and accelerates the detection of persons infected with coronavirus as well as monitors their movements to reduce disease spread. Bhatia et al. [24] have employed a Multi-scaled Long Short Term Memory (M-LSTM)-based vulnerability prediction for preventive veterinary healthcare services. Moreover, a fog-assisted real-time alert generation module has been presented in the authors' framework to notify the concerned veterinary doctor in the case of medical emergency.

The overview of existing approaches for the development of multi-layered architectures for smart systems shows that the intelligent and coordinated management of the three-layer IoTFC model has been the subject of many studies. In such an ecosystem, if a task requires a large amount of computing resources or data storage space, the processing should be performed in the cloud layer. In the case of a task needing low latency, the sensors and devices should send the data to servers or computing devices in the fog layer.

2.2. Literature Review on Design Peculiarities of Smart Home Systems with NLU Interface

Many authors have developed smart home systems with voice recognition interface using classical and machine learning approaches. Intent classification from utterances is typically performed in a two-stage pipeline: first, extraction of transcriptions using Automatic Speech Recognition (ASR) and second, intent recognition. In order to reduce the errors accumulated at the ASR stage and their negative impact on the intent classifier, Desot et al. [25] have proposed an end-to-end model based on deep neural networks to perform intent classification directly from raw speech input. Liu et al. [26] have applied deep learning model to multi-intent detection task. Klaib et al. [27] have used Amazon

Alexa in the new smart home system to a receive user's verbal commands and then send the request to Azure cloud to control home appliances. Yang [28] has designed and implemented a new voice recognition smart home system with client-server architecture. The intelligent terminal (client) interprets the voice signal as a specific voice command using a hidden Markov model and a dynamic time warping algorithm. The control node finds the system's command corresponding to the voice command, sends it to the home devices, and responds to the system's control panel. Amin [29] has presented a smart home system as a single portable unit that uses a voice-controlled Android application to operate home devices. The new system utilises ThingSpeak cloud platform to collect, review and store data from home appliances. A web server retrieves data from ThingSpeak cloud and saves it into MySQL database. The new home automation system applies a Google voice recognition service and a microcontroller to execute certain voice commands. Stefanenko et al. [30] described a method for voice commands recognition based on fuzzy logic. The developed fuzzy system has been employed to execute linguistically inaccurate commands. The obtained results show that the proposed method increases the expressiveness of the voice control of a moving robot.

In the above-mentioned studies, the positive and negative characteristics of VUI have been identified and various applications of VUI based on different speech recognition methods have been presented. After the analysis of the literature review for the development of intelligent voice interface, the following conclusions can be drawn:

1. Voice user interface can be implemented only by integration of methods covering all aspects of natural language processing (voice data entry, tokenisation, lemmatisation, tagging, semantic analysis);
2. Intent detection methods can be categorised into three main groups:
    - Methods using statistical features (hidden Markov model, dynamic time warping, naive Bayes, AdaBoost, support vector machines, logistic regression) [28];
    - Neural networks (convolutional neural networks, recurrent neural networks) and deep learning (LSTM), distance-based (Term Frequency-Inverse Document Frequency–TF-IDF) methods or combination of several deep learning methods [25,26];
    - Other intelligent methods for semantic recognition of voice commands (fuzzy logic [30], semantic patterns).
3. Since advanced-level voice recognition devices are costly for mainstream household appliances, the developers preferred commercial speech recognition services, such as Amazon Alexa [27], Google Assistant [29].
4. The implemented interactive user interfaces are only focused in a specific subject area [25,27,28].

In summary, the authors of the above-mentioned studies have employed statistical and machine learning methods for voice command classification in the most widely used natural languages. The first issue of existing approaches is the limited number of supported natural languages. There are several commercial NLU platforms available in the market at affordable prices such as Alexa Skills (2015) [31], Azure Cognitive Services (2018) [32], Watson Assistant (2020) [33] and Dialogflow CX (2021) [34]. Despite the fact that these platforms offer general-purpose natural language services with a lot of functionality, they have some disadvantages: (1) Their NLU services only support several languages (a maximum of 13 to 87). Bulgarian and many other under-resourced languages are not supported. (2) Natural language processing services are quite complex to test, set up and maintain. Further challenges are as follows: time delays due to remote data processing in the cloud, increased costs (as voice services are paid), and privacy and security issues due the risk of personal data breach.

In order to achieve the goal of our research, it is necessary to develop a new NLU service for speech recognition of domain-specific commands in an under-resourced language. For this purpose:

1. Each input user's utterance needs to be converted to its semantic equivalent (intent).
2. Specific syntax for encoding structured semantic templates must be defined for smart home commands (nodes, patterns and slots);
3. Speech-To-Text and Text-To-Speech services offered by Microsoft, Google and other speech service providers are among the building blocks of the new intelligent voice user interface. In the case of normal speed and reliable Internet connection, they could be employed via the cloud or otherwise as on-premises software.

The same approach can be applied for any language with available STT and TTS services and not supported intent recognition.

3. New Framework for Voice-Based User Interface

In order to achieve interconnection and interoperability of multiple devices, services and applications in intelligent home systems, there are two main requirements for their architectural design.

1. Smart home systems have to provide intelligent user interface, aimed at maximum user convenience.

Intelligent user interface can personalise and guide interaction. IUI is one of the most important and distinguishing characteristics that determines, to a large extent, new products adoption, especially in regional markets. Voice communication options improve customer experience, increase consumer loyalty and create a competitive advantage for new products. It also enables initial installation and configuration, periodic diagnostics and maintenance, reconfiguration and optimisation of smart systems.

2. Home automation systems should guarantee effective remote communication between the control panel and other control systems (IoT networks, building management system, etc.).

In order to meet this second requirement, data has to be transferred via a gateway. In this case, the control panel of a smart home system should be transformed into a hub, supporting standardised IoT communication protocols, such as WiFi and KNX, a communication protocol developed for and widely used in home automation. The voice control of smart home systems could be implemented through an Alexa Skills mechanism or similar.

According to the two aforementioned aspects of smart home connectivity, the new framework needs to combine an innovative user interface and diverse communication capabilities between the control panel and other (external) control systems.

As pointed out in the previous section, the voice user interface can be implemented in two ways:

- The functionality is embedded in the control panel of the smart home systems, which has a secured a connection to the cloud infrastructure;
- The functionality is located in a separate layer within the structure of smart home systems.

The new conceptual framework for smart home systems with VUI follows the second approach (Figure 1).

The diagram in Figure 1 visualizes the communications of smart home components with system's panel via two types of channels: (1) direct: to the peripheral devices and external computer networks and mobile networks; and (2) indirect: to IP sensors, devices and appliances, other control systems (KNX) and voice-based interface.

In the proposed framework, an integration server (a fog structure) is introduced as an additional communication node. It allows for both the deployment of the intelligent user interface as well as integration with existing IoT devices and control systems.

Another novel element in the proposed framework is that communications between users (left) and external control systems (right) can be realised via virtual channels to a variety of services offered by different providers, including by smart automation company.

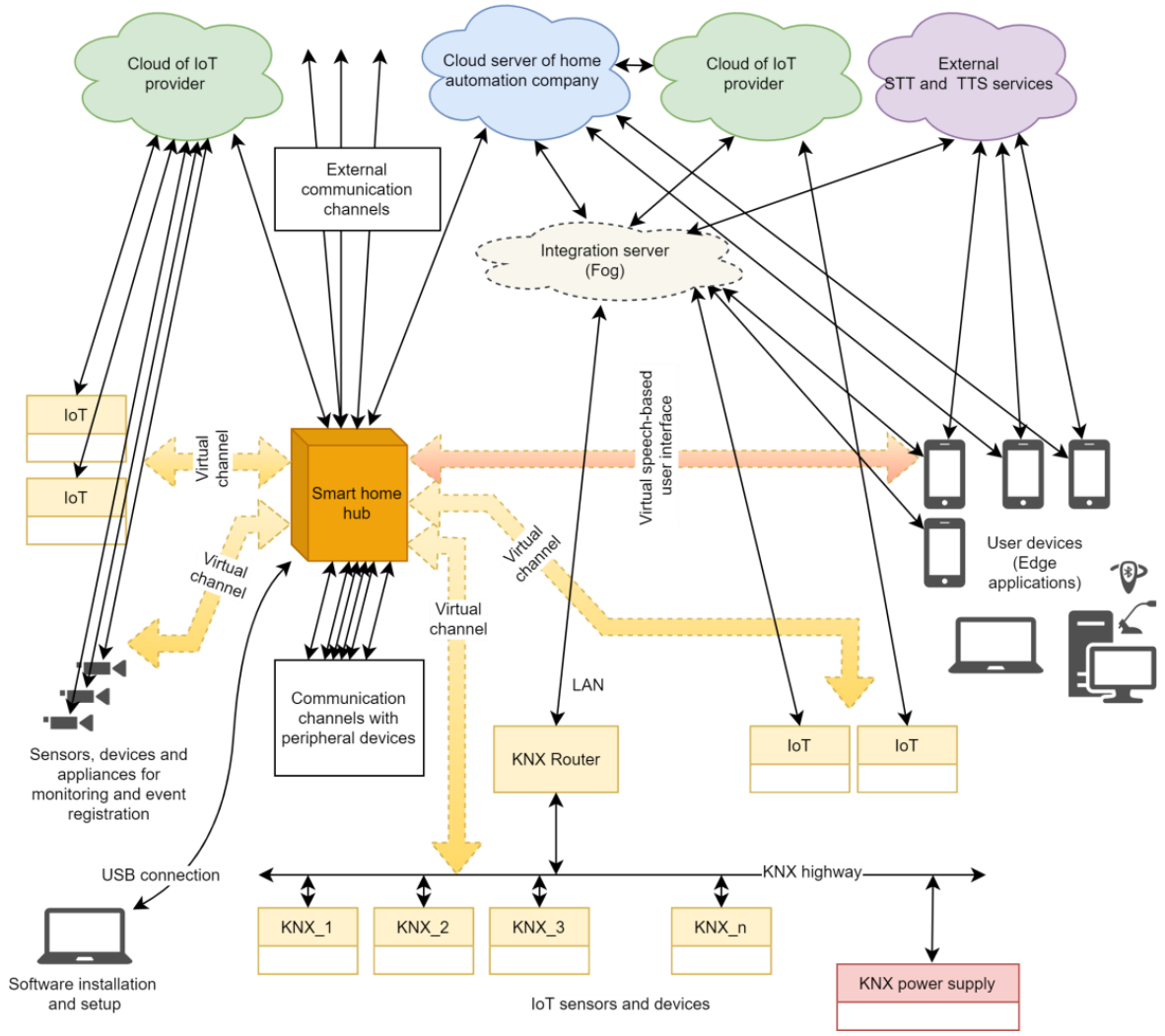


Figure 1. Framework for voice-based user interface. Note: Black arrows indicate direct communications between central panel (hub) and system's elements, while yellow arrows show indirect ones. The indirect communications consist of multiple requests to services.

The new software module implementing a dialog-based IUI can be deployed in any layer of the smart home architecture. Placing the voice interface in a separate architectural level offloads the control panel from non-core functionality and distributes computations between different devices. This approach is preferred because:

- IUI development is relatively independent from system panel limitations;
- The risks of security breaches are shifted away from the system core.

Another advantage of this approach is that multiplier effect can be utilised by implementing the same intelligent interface module for an entire product family and in different languages.

The proposed distributed multi-layer framework expands the capabilities of smart home systems in comparison with classical IoTFC approach. Edge or fog voice services can provide quick responses to requests with minimal delays because processing is located where data is available or needed.

4. System Prototype Design, Implementation and Evaluation

This section presents the implementation of a smart home system in line with the proposed framework for voice control combining local and remote natural language services for under-resourced languages with visual illustrations and algorithms' descriptions.

4.1. Design of Interactions between User and a Smart Home System

The flowchart in Figure 2a and corresponding steps in Algorithm 1 represent the voice-based interactions between a user and a smart home system as a sequence of requests and responses.

Algorithm 1 User–smart home system interactions via speech command as a part of SUI

1. Smart home user enters an utterance (speech) with a certain meaning (a1);
2. The user device (computer, laptop, tablet, smartphone or wearable) records the speech and transmits it as an audio stream (a2) to a service provided by the integration server;
3. IUI software module receives the message. The module sends a request (a3) to STT cloud service, forwarding the audio data;
4. IUI receives a response (a4) to its request. It is a text that corresponds to the sent audio data.
5. IUI extracts the meaning of the input user utterance and generates a request containing the obtained command for the smart home hub. The request (a5) is sent to the server in Cloud of Home Automation Company (CHAC);
6. CHAC receives the request, verifies its authenticity and correctness, prepares and sends the command (a6) to the smart home sub-system specified for it;
7. The smart home sub-system receives the command, executes it and returns the corresponding response (a7);
8. The response is received by CHAC, which transforms it to an answer to an IUI re-quest and sends this answer (a8) to IUI;
9. IUI receives the response from CHAC, prepares its own text response and sends it through a request (a9) for conversion to TTS cloud service;
10. IUI receives the result of its request as an audio file (a10);
11. IUI prepares a response specifying the URI of the received audio file and sends it (a11) to the user device;
12. The user device plays the audio file (a12) received as a result of their request and the user hears a speech response by the system to their utterance.

In some cases, several steps of Algorithm 1 can be omitted and/or added (Figure 2b):

1. Audio data from the user device can be sent directly to STT service (skipping step a2), and the received response is then submitted to IUI service (step b4 is added).
2. During the preparation of a response as an audio file, the call to TTS service might be skipped (steps a9 and a10 can be omitted) if such response has already been made and catalogued on the integration server. In this case, the URI to the already created audio file is sent as a response to the user device.

The flowchart in Figure 2b visualises the shortest version of this interaction process. In this way, the communication between user and smart home system takes less time and is more efficient.

The proposed Algorithm 1 is composed according to the classical five-stage structure of voice-controlled dialogue [35] (Table 1).

Table 1. Description of voice-controlled dialog in a smart home system (in two versions; a more detailed and the shortest one).

Stage	Detailed Version	The Shortest Version
Voice Activity Detection	a1	b1
ASR	a2–a4, a9–a10	b2–b3
NLU	a5	b5
Decision Stage	a6–a7	b6–b7
Communication Stage	a8, a11–a12	b8, b9–b10

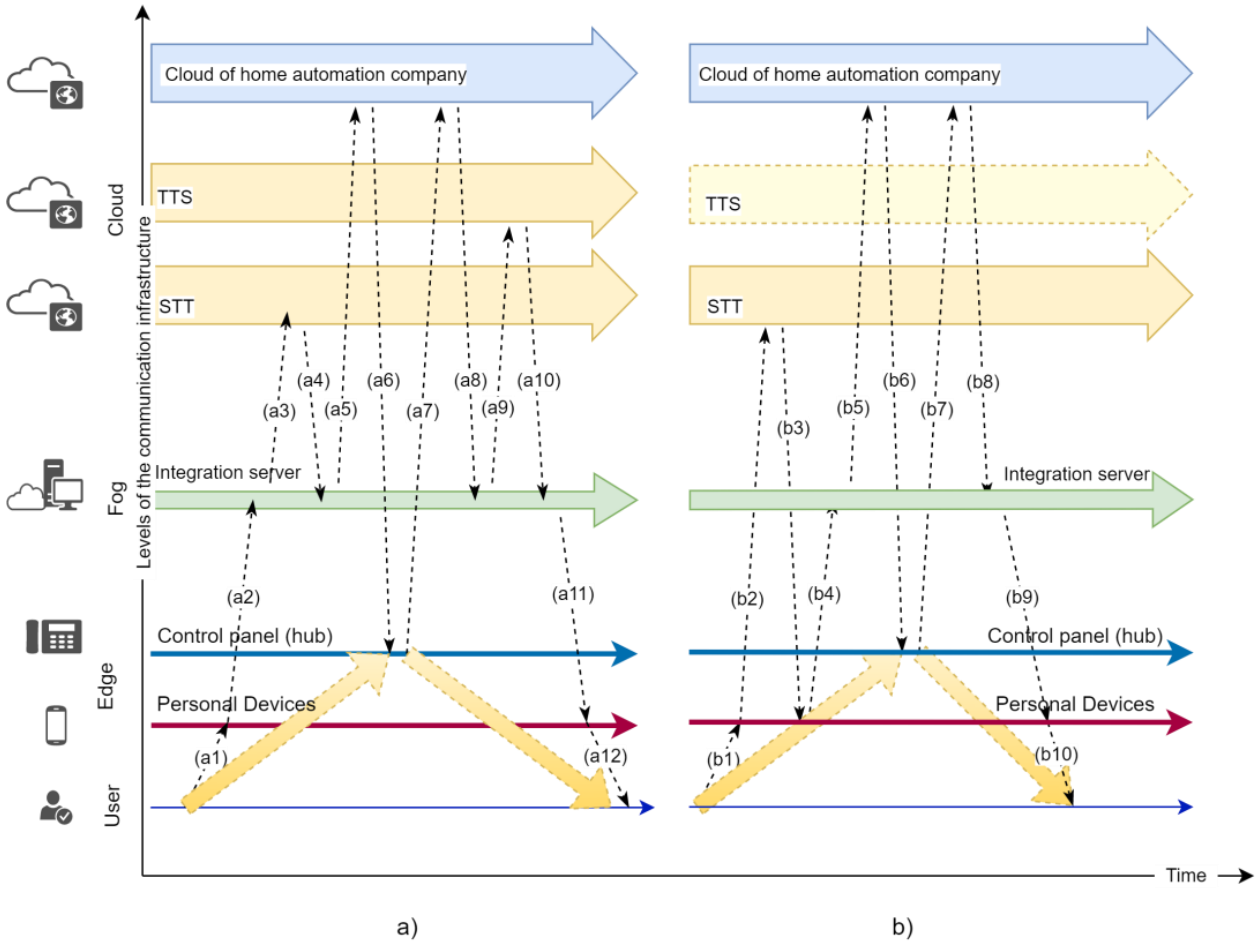


Figure 2. Step-by-step voice communication between user and smart home system, (**a**) long and (**b**) short version, respectively.

The flowchart in Figure 3 and corresponding steps in Algorithm 2 depict the proposed intent detection method (Algorithm 1, Step 5). In the algorithm description, a node represents a collection of predefined information structures, which is necessary to recognise each possible executable command. They are stored as a list of nodes. Each node contains a list of structural patterns, a list of slots, and command generation mechanism.

Each command structural pattern is defined by a specialised syntax and contains word lemmas, synonym groups, parameter slots, etc. It defines the permissible positions of keywords and parameter slots in the utterance. It is used as a basis for detecting the correspondence between the utterance and the pattern.

Each slot corresponds to one parameter required by the command. Specific questions for requesting the value of the corresponding parameter from the user are defined for it.

Command generation mechanism is a procedure in which the received command parameter values are processed and a formal command is generated for the hub. It can include other procedures, links to external services and URL references.

The list of commands can be easily extended with no coding required.

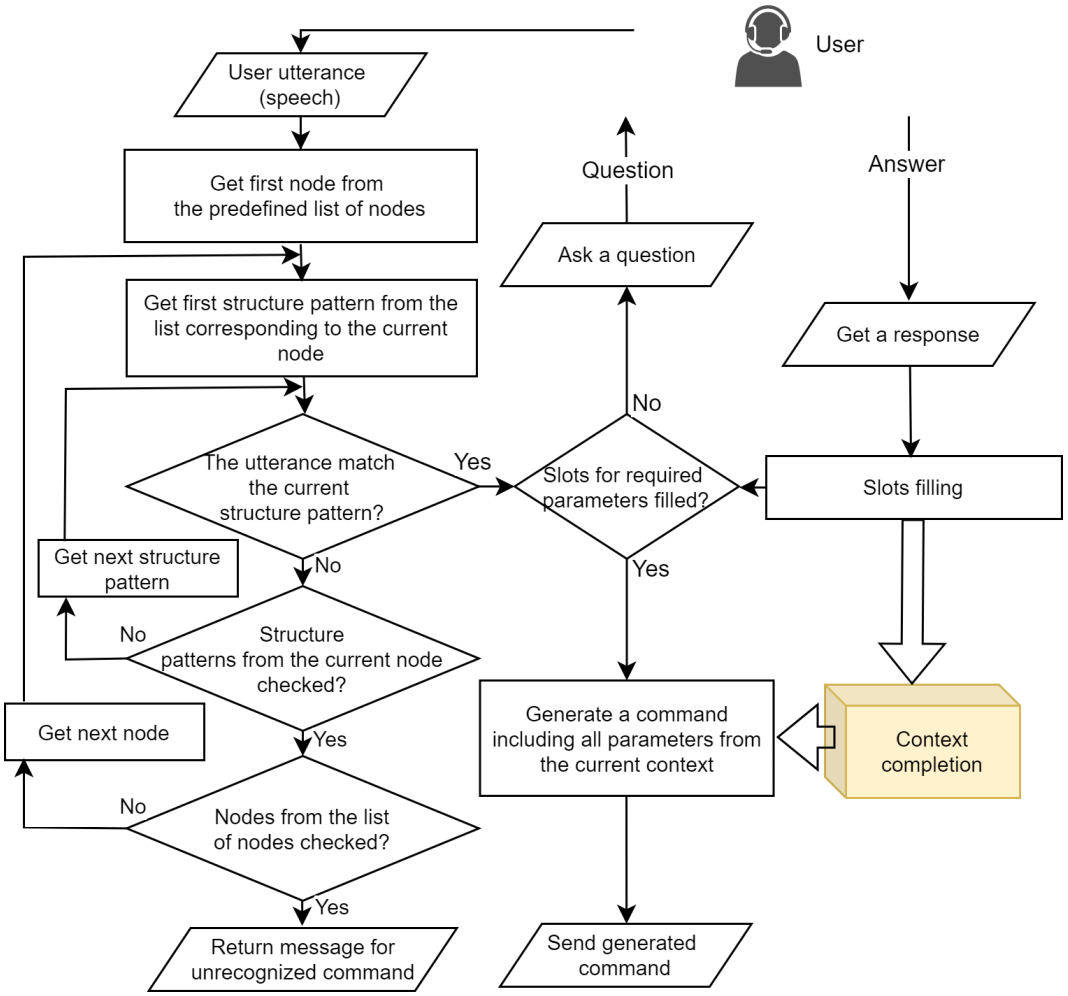


Figure 3. Flowchart of intent-based voice command recognition.

Algorithm 2 Intent-based voice command detection from user utterance

1. User utterance is entered;
2. All nodes from the list of predefined application nodes are traversed sequentially. The process starts with the first one in the list;
3. For each node, all structural patterns associated with this node are traversed sequentially;
4. The utterance is compared with the current structural pattern;
5. If there is a match, the traversal of the nodes stops and the command corresponding to the current node is considered as recognised;
6. If there is a recognised command, it is checked whether all required parameters are available in the utterance or in the current context;

6a. If the context still does not contain values for all required parameters, the user is asked clarifying questions about each missing value and upon receiving the answers, the values are filled in the context;

6b. If all required parameters have corresponding values in the context, a command is generated, including the parameter values accumulated in the context, and it is sent to the execution device;

7. In the case of missing match after traversing of all nodes and all their structural patterns, an unrecognised command message is generated and sent to the user.

In the next subsection, we present the peculiarities of software implementation of new VUI for existing smart security system and the results obtained from demonstration of new system prototypes in an operational environment.

4.2. Implementation Details and Evaluation

In order to validate the functionality of the new voice user interface, the existing smart security system has been extended with additional embedded software and software modules on the server as follows:

1. System software for integration server has been developed.

The integration server has been positioned in the intermediate (fog) layer of the proposed IoTFC framework. On this server, new software for voice command interface in domain-specific Bulgarian has been deployed. The new user interface has been implemented as a Web application. The interface has been connected to STT and TTS cloud services (both in Bulgarian). The program code has been created in Node.js environment;

2. Several additional services have been added to the existing cloud of smart home automation company;
3. New services have been created to handle remote users' requests for control and monitoring of the smart home system including from third party applications;
4. An add-on has also been created to employ Alexa Skills and other similar applications for mobile access and control of system devices.

The main challenge during the development process of our smart voice interface was intent extraction from the users' utterances and its conversion into a specific structured command. The issues that must be overcome are as follows:

1. The computing power of personal and mobile devices is limited;
2. The computational complexity of STT, TTS and natural language understanding tasks is significant.

The implemented algorithm also employs a deep learning neural network, which is pre-trained to recognise wake-up words through TensorFlow platform (in its JavaScript version), installed on user devices. This approach eliminates redundant traffic that would occur if user devices "listen" to any noise and forward it for recognition somewhere in the network. The rest of the natural language understanding (Algorithm 2) is performed on the server side. In this way, personal devices are unloaded from some tasks and the know-how of manufacturer of smart security system is protected.

After the implementation of VUI for smart security system and its connections to local and remote services, a demonstration of system prototypes in operational environment was organised. Two different prototypes of wireless smart security system were presented. They were configured to work with various wireless devices: passive infrared (PIR) sensors in different zones, magnetic sensors, external siren, remote control devices, etc. During the demonstration using voice control in Bulgarian, the following system components and features were tested:

- Wake-up of listening modules with a predefined word;
- Activation (turn on) of security system in arm mode; an indication has been received on the panel display that arm mode is turned on;
- Lights control in different rooms, with visual and verbal confirmation of recognised and executed commands;
- Hints are provided for command options when a dialogue has already been started but some slots are still empty;
- Providing information about the projects (distribution of projects information);
- Speech commands responses by two synthetic voice types (male and female);
- Several command variations have been executed with different keywords, keywords' derivatives and paraphrases and the extracted intent has been validated.

New services based on the proposed framework for VUI have been built into an existing smart security system and the systems' prototypes has been demonstrated in an operative environment. The test results have shown that the system can successfully control security facilities through a predefined set of voice commands, taking into account real-time environmental data. The experimental results also demonstrate that the new VUI is reliable and flexible; it can execute speech commands with different variations in users' utterances and voice nuances.

In this section, an intelligent home system, which employs the proposed IoTFC framework, has been presented. The new system can react immediately to all events that are signalled, i.e., the system operates normally in isolated (local) mode. This smart home system can collect and process data generated by its peripheral (edge) devices (sensors, devices and appliances) and control them remotely in real time. It can also communicate with user's smartphones and wearable devices, but in these cases, Bluetooth communication protocol is required. The proposed architecture for smart home allows for integration with other external systems (for example, systems for building automation), acting as a hub connecting different systems. The smart system can transmit IoT data to the fog layer via REST HTTP protocol, which provides flexibility and interoperability in building RESTful services. This feature ensures backward compatibility with legacy systems running in the computational infrastructure of smart home automation company. Local computing entities (fog nodes) can filter received data and either process it locally or send it to the cloud for further processing. The voice user interface provides control and monitoring services using real-time data about devices' state and events occurring in the home environment. It can be deployed in any layer of the smart home architecture. The voice user interface can be implemented in any non-tonal language, including under-resourced ones, if STT and TTS services are available for it.

5. Conclusions

Edge-Fog-Cloud computing and automatic speech recognition are among the most dynamic areas of today's computer science. In our manuscript, we combine IoT-Fog-Cloud architecture and speech recognition methods to develop a new distributed framework for smart home systems with voice user interface for under-resourced languages. The new framework incorporates existing STT and TTS cloud services in speech recognition method in a particular (smart home) domain.

The limitations of our study are as follows: (1) Defining patterns for smart home commands requires very good knowledge in the respective language(s). For example, since Bulgarian and the other Slavic languages belong to the group of morphologically rich languages, special attention is needed to different word forms (morphological derivation) in patterns of smart home commands. (2) The prototype implementation can only handle a predefined set of domain-specific voice commands; for smart security system, in only one language (Bulgarian). (3) The proposed methodology does not support VUI in tonal languages and in languages without available STT and TTS services. (4) The commands happen only at the sentence level.

Our future research directions are as follows: (1) implement the proposed framework using small single-board computers; (2) modify the proposed framework for voice command interface for industrial automation systems; (3) enhance the prototype by adding more voice commands for monitoring and controlling of smart home devices and appliances in several natural languages. In the future, we also plan to apply fuzzy multi-criteria decision-making methods in voice-controlled human-machine dialogues.

Author Contributions: Conceptualisation, Y.I.; software, Y.I.; validation, G.I. and Y.I.; resources, G.I. and Y.I.; writing—original draft preparation, G.I.; writing—review and editing, G.I. and Y.I.; visualization, G.I and Y.I.; supervision, G.I.; project administration, Y.I.; funding acquisition, G.I. and Y.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Innovation Fund, Grant No. 11IF-02-5/01.12.2020 "Scientific and applied research for the development of a wireless security system with universal connectivity" and by the Ministry of Education and Science and by the National Science Fund, co-founded by the European Regional Development Fund, Grant No. BG05M2OP001-1.002-0002 "Digitization of the Economy in Big Data Environment".

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the academic editors and anonymous reviewers for their insightful comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schiefer, M. Smart home definition and security threats. In Proceedings of the 2015 Ninth international conference on IT security incident management & IT forensics, Magdeburg, Germany, 18–20 May 2015. [CrossRef]
2. Domb, M. Smart home systems based on internet of things. In *Internet of Things (IoT) for Automated and Smart Applications*; Ismail, Y., Ed.; IntechOpen: London, UK, 2019; pp. 25–40. [CrossRef]
3. Stojkoska, B.L.R.; Trivodaliev, K.V. A review of Internet of Things for smart home: Challenges and solutions. *J. Clean. Prod.* **2017**, *140*, 1454–1464. [CrossRef]
4. Wei, Z.; Qin, S.; Jia, D.; Yang, Y. Research and design of cloud architecture for smart home. In Proceedings of the 2010 IEEE International Conference on Software Engineering and Service Sciences, Beijing, China, 16–18 July 2010. [CrossRef]
5. Soliman, M.; Abiodun, T.; Hamouda, T.; Zhou, J.; Lung, C.H. Smart home: Integrating internet of things with web services and cloud computing. In Proceedings of the 2013 IEEE 5th International Conference on Cloud Computing Technology and Science, Bristol, UK, 2–5 December 2013. [CrossRef]
6. Wadhwa, H.; Aron, R. Fog computing with the integration of internet of things: Architecture, applications and future directions. In Proceedings of the 2018 IEEE International Conference on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications, Melbourne, Australia, 11–13 December 2018. [CrossRef]
7. Atlam, H.F.; Walters, R.J.; Wills, G.B. Fog Computing and the Internet of Things: A Review. *Big Data Cogn. Comput.* **2018**, *2*, 10. [CrossRef]
8. Rahimi, M.; Songhorabadi, M.; Kashani, M.H. Fog-based smart homes: A systematic review. *J. Netw. Comput. Appl.* **2020**, *153*, 102531. [CrossRef]
9. Shukla, S.; Hassan, M.F.; Khan, M.K.; Jung, L.T.; Awang, A. An analytical model to minimize the latency in healthcare internet-of-things in fog computing environment. *PLoS ONE* **2019**, *14*, e0224934. [CrossRef] [PubMed]
10. La, Q.D.; Ngo, M.V.; Dinh, T.Q.; Quek, T.Q.; Shin, H. Enabling intelligence in fog computing to achieve energy and latency reduction. *Digit. Commun. Netw.* **2019**, *5*, 3–9. [CrossRef]
11. Worldwide Smart Home Devices Market Grew 11.7% in 2021 with Double-Digit Growth Forecast Through 2026. According to IDC. Available online: https://www.idc.com/getdoc.jsp?containerId=prUS49051622 (accessed on 30 November 2022).
12. Smart Home Devices to Exceed 13 Billion in Active Use by 2025, with Entertainment Devices Leading Way. Available online: https://www.juniperresearch.com/press/smart-home-devices-to-exceed-13-billion-in-active (accessed on 30 November 2022).
13. Sovacool, B.K.; Del Rio, D.D.F. Smart home technologies in Europe: A critical review of concepts, benefits, risks and policies. *Renew. Sust. Energ. Rev.* **2020**, *120*, 109663. [CrossRef]
14. DialogFlow. Available online: https://cloud.google.com/dialogflow/ (accessed on 30 November 2022).
15. Sabharwal, N.; Agrawal, A. Introduction to Google Dialogflow. In *Cognitive Virtual Assistants Using Google Dialogflow*; Apress: Berkeley, CA, USA, 2020; pp. 13–54.
16. What is Language Understanding (LUIS)? Available online: https://learn.microsoft.com/en-us/azure/cognitive-services/luis/what-is-luis (accessed on 30 November 2022).
17. Rozga, S. Language Understanding Intelligent Service (LUIS). In *Practical Bot Development*; Apress: Berkeley, CA, USA, 2018; pp. 47–128.
18. Watson Natural Language Understanding. Available online: https://www.ibm.com/cloud/watson-natural-language-understanding (accessed on 30 November 2022).
19. Vergara, S.; El-Khouly, M.; El Tantawi, M.; Marla, S.; Sri, L. *Building Cognitive Applications with IBM Watson Services: Volume 7 Natural Language Understanding*; IBM Redbooks: Armonk, NY, USA, 2017; pp. 1–57.
20. Alibaba Cloud Intelligent Speech Interaction. Available online: https://www.alibabacloud.com/product/intelligent-speech-interaction (accessed on 30 November 2022).
21. Sun, H.; Yu, H.; Fan, G.; Chen, L. Energy and time efficient task offloading and resource allocation on the generic IoT-fog-cloud architecture. *Peer Peer Netw. Appl.* **2020**, *13*, 548–563. [CrossRef]
22. Chegini, H.; Naha, R.K.; Mahanti, A.; Thulasiraman, P. Process automation in an IoT–fog–cloud ecosystem: A survey and taxonomy. *IoT* **2021**, *2*, 6. [CrossRef]
23. Kallel, A.; Rekik, M.; Khemakhem, M. IoT-fog-cloud based architecture for smart systems: Prototypes of autism and COVID-19 monitoring systems. *Softw. Pract. Exp.* **2021**, *51*, 91–116. [CrossRef]

24. Bhatia, M.; Sood, S.K.; Manocha, A. Fog-inspired smart home environment for domestic animal healthcare. *Comput. Commun.* **2020**, *160*, 521–533. [CrossRef]
25. Desot, T.; Portet, F.; Vacher, M. Towards end-to-end spoken intent recognition in smart home. In Proceedings of the 2019 International Conference on Speech Technology and Human-Computer Dialogue, Timisoara, Romania, 10–12 October 2019.
26. Liu, J.; Li, Y.; Lin, M. Review of intent detection methods in the human-machine dialogue system. *J. Phys. Conf. Ser.* **2019**, *1267*, 012059. [CrossRef]
27. Klaib, A.F.; Alsrehin, N.O.; Melhem, W.Y.; Bashtawi, H.O. IoT Smart Home Using Eye Tracking and Voice Interfaces for Elderly and Special Needs People. *J. Commun.* **2019**, *14*, 614–621. [CrossRef]
28. Yang, C. Design of smart home control system based on wireless voice sensor. *J. Sens.* **2021**, *2021*, 8254478. [CrossRef]
29. Amin, D.H.M. Voice Controlled Home Automation System. *Int. J. Electr. Comput. Eng.* **2022**, *7*, 1–11. [CrossRef]
30. Stefanenko, O.S.; Lipinskiy, L.V.; Polyakova, A.S.; Khudonogova, J.A.; Semenkin, E.S. An intelligent voice recognition system based on fuzzy logic and the bag-of-words technique. *IOP Conf. Ser.: Mater. Sci. Eng.* **2020**, *1230*, 012020. [CrossRef]
31. Alexa Skills. Available online: https://developer.amazon.com/en-US/alexa/alexa-skills-kit/ (accessed on 30 November 2022).
32. Azure Cognitive Services. Available online: https://azure.microsoft.com/en-us/products/cognitive-services/ (accessed on 30 November 2022).
33. Watson Assistant. Available online: https://www.ibm.com/products/watson-assistant (accessed on 30 November 2022).
34. Dialogflow CX. Available online: https://cloud.google.com/dialogflow/cx/docs/ (accessed on 30 November 2022).
35. Vacher, M.; Caffiau, S.; Portet, F.; Meillon, B.; Roux, C.; Elias, E.; Lecouteux, B.; Chahuara, P. Evaluation of a Context-Aware voice interface for ambient assisted living: Qualitative user study vs. quantitative system evaluation. *ACM Trans. Access. Comput.* **2015**, *7*, 1–36. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.