CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

Informatics

Department of Information Technologies

Bachelor thesis

Voice control in mobile applications

Author: Gayrat Dadamirzaev

Supervisor: Ing. Miloš Ulman, Ph.D.

Declaration

I hereby declare that I have worked on this bachelor thesis, named "Voice control in mobile applications", myself and only the materials listed at the end of this paper has been directly cited.

In Prague,…….. ……

Gayrat Dadamirzaev

Acknowledgment

I would like to express and direct gratitude to Ing. Milos Ulman for his strong and significant recommendations and advices, coordination and supervision of corresponding bachelor thesis, to represent brightly idea of the topic.

Along with this, I thank my group mates of university for discussing and for sharing ideas to make this thesis more successful.

Voice control in mobile applications

Hlasové ovládání v mobilních aplikaci

Voice control in mobile applications

Summary

The current bachelor thesis is focused on the topic of the Voice control in mobile applications. The first part of the thesis is the theoretical framework, which defines the history of appearing of Artificial Intelligence (AI). The following information is provided: descriptions of the development and use of AI, Turing test, etc. The second part covers the knowledge about the speech recognition system in machine, such as the voice command process and specific features of industrial voice control system. This part consists of the theoretical specification of speech control, the description of first mobile application and voice control, comparison leaders of voice control in mobile application such as Samsung S-voice, Apple Siri, Windows Voice and Blackberry Voice. In addition, it is given an overview of each application.
The third part is presented in the form of the analytical and practical research. The analysis are concentrated on the evaluation of each application, which includes Samsung, Apple, Windows and Blackberry smartphones by using the Multi-Criteria Analysis for decision making. This ranking method represents the best and optimal decision in accordance to the theoretical part, own solution and questionnaire of three independent experts. The results are analyzed according to the certain criteria and lead to the determination of the best voice control.

Keywords

AI(Artificial Intelligence), Turing test, voice command process, industrial voice control system, voice search, Google Voice, Siri

Hlasove ovladani v mobilnich aplikaci

Souhrn

Tenta bakalařská prace je zaměrena na temata Voice control in mobile applications. První část prace obsahuje teoritickou báze vysvětlující historie vytvoření Úmělé Inteligence (UI). Nasledující informace vymezuje pojmy: popsaní vzníku a použití UI, Turing test, atd. Druhí čast prace informuje o systemu pozpoznávání řečí ve zdrojích, tak jako proces hlasového ovladání a specifické rysy průmyslového systemu hlasové kontroly. Tato část obsahuje teoretické pojmy hlasového ovladání, popsání mobilní aplikace, jako Samsung S-voice, Apple Siri, Windows Voice and Blackberry Voice. Pro doplnění je přepraveno definování každý aplikace.

Třetí část ja vytvořena v podobě analetického a praktického výzkumu. Analýzy zaleží na ohodnocení každé aplikace, například Apple, Windows and Blackberry smartphones, využívající Multi-Criteria Analysis pro rozhodování. Tato metoda přezentuje nejlepší a optimalní řešení na zakladě teoretické častí osobní řešení a dotazník třech nezávislých expertů. Vysledky jsou ohodnoceny kvůli určítých měřítcích a vmezují nejlepší rozhodování.

Klicova slova

Umělá intelligence, Turingův test, proces hlasového příkazu, průmyslové řídící hlasových system, hlasové vyhledávání, Google Voice, Siri.

Table of Contents

1. Introduction

Day by day humankind making a new step in progress, create a new idea, new technology, with purpose making easier and faster task for maximum saving time and energy of the people.

Not a long time ago, we were not able to think that today we can chat with each other, see each other via devices, using pc and mobile phones. However, today, conversely, people are so adapted and dependent of the virtuality, that nowadays it is hard to imagine people's regular life without this equipment, which are surrounding men everywhere.

The scientific name of human is homo sapiens-man the wise- because people's mental capacities are so important to everyday lives and the sense of self. The field of artificial intelligence, or AI, attempts to understand intelligent entities.AI has produced many significant and impressive products even at this early stage in its development. Although no one can predict the future in detail, it is clear that computers with human-level intelligence (or better) would have a huge impact on our everyday lives and on the future course of civilization.

AI is one of the newest disciplines. It was formally initiated in 1956, when the name was coined, although at that point work had been under way for about five years.

AI currently encompasses a huge variety of subfields, from general-purpose areas such as perception and logical reasoning, to specific tasks such as playing chess, providing mathematical theorems, writing poetry, and diagnosing diseases. Often, scientists in other fields move gradually into artificial intelligence, where they find the tools and vocabulary to systemize and automate the intellectual tasks on which they have been working all their lives.

In the field of building intelligent mobile assistance systems, AI is not only a nice add-on but also a necessary pre-requisite. Moreover, not just one AI technique is needed here. As we can see in the contributions to this workshop, a whole set of different techniques are to integrated in order to reach a satisfying goal. Learning, for instance, is crucial for user-adapted services. Both symbolic AI learning methods and sub-symbolic AI methods from the field of neural networks can provide the necessary flexibility for a system that adapts to the user's preferences in different contexts. In addition, spatial

reasoning becomes a central component of location aware systems. Here, these methods are challenged with real world environments and it can be shown that they are not only useful in some toy-environment of a block world simulation. Other techniques may be used for resource adaptive services and for smart localization of the user, e.g. through decision-theoretic planning of navigation instructions. (RUSSELL, Stuart J a Peter NORVIG. *Artificial intelligence: a modern approach*. Englewood Cliffs, N.J.: Prentice Hall. ISBN 01-310-3805-2.)

This new workshop Artificial Intelligence in Mobile Systems (AIMS) brings together researchers working in the sub-fields of AI described above and those working with the design of mobile applications and devices. The workshop provides a platform for presentations and discussions on these various topics related to mobile systems.

Thanks to this achievement and tests now we are able to create and use Voice Control or Voice Recognition in operational systems, mobile applications, such as Android, Symbian, iPhone, Windows Mobile, BlackBerry, Java Mobile Edition, Linux Mobile.

2. Thesis objective and methodology

2.1. Objective

Exploring the topic Voice control in mobile application, there were delivered the main goals for widely studying certain aspects directly related to topic such as human progress before and now. In the first part or the first objective is to describe an overview of artificial intelligence, its importance and necessity, relevance and actuality, demonstrate a brief history in which there is written about the Turing's test.

Approaching to the next objective, author proceeding to the second part of thesis. The second goal is a principle part, analysis and stepwise description of practical part. The phased description will be given a real example of voice control, such as Google voice, Siri from Apple. In this case, the author selects the some voice recognition programs on several types of mobile devices, such as Android, iOS, Symbian etc. There is clearly displayed functionality, properties of each applications. The program is designed to user that could easily extend to using with application, and the application function must work properly with the minimum amount of mistakes. The integration between user and devise is when user speaks some name from contact list of his mobile, it automatically should request for dialing to the owner of the number. This part consists of feedback and discussions about the usefulness, advantages of voice recognition systems.

The partial goal is about criteria for evaluation, for which author is using special method, which calls Multi-Criteria Analysis of variances (MCA). Based on information from the second part and following to it, it could be possible to determine which one of the program

works best and which one is the worst. This part will show the significant idea of the thesis, which is based to the objectives below.

2.2 Methodology

Source of materials, information, facts and hypotheses were gathered by the author, using research articles, book and newspaper publishing..

The basis of the methodology is the first step, which will be taken into consideration almost all aspects of voice command processing (speech recognition, syntactic and semantic analysis and spontaneous speech effects) and - most importantly - their mutual influence. In the same article, it would be examined the following methods:

Assay of the literature related to artificial intelligence and mobile applications; Multi-Criteria Analysis of decision making ,Comparison between leaders of smartphones market Apple and Samsung; synthesis; Theoretical analysis and synthesis; abstraction; induction and deduction; classification; generalization;

The second part will be examined using the investigation method that is called MCA (Multi-Criteria Analysis). In assuring method will check the practice analysis, its gradual steps, decision, the use of investigation, and the most important one is results. Based on the theoretical knowledge and practical example the conclusion will be formulated.

3. Artificial Intelligence.

John McCarthy

The Artificial Intelligence (AI) is a science, creations of intelligent machine that carry out tasks, and better and quicker, than people can. AI actually is not about intelligence, though. It is a question of a solution. Nevertheless, decisions in the field of artificial intelligence always are mathematics and computer.

Science birth.

Until 1950th years, the phrase in itself does not exist in general! However, all this changed, thanks to John McCarthy, "*The Father of Artificial Intelligence*".

John McCarthy is mathematician who considers that machines can be made the reason, as well as people. In 1956, it sponsored the Dartmouth-year Research Project on AI to investigate this opportunity. At conference, McCarthy collected some of the most creative thinkers in the field of mathematics, logic, engineering, psychology, and then new area of informatics. As a result of this conference, artificial intelligence became recognized as separate area of research.

3.1 Ideas and citations of scientists.

Previously, when the idea of creating intelligent computer came to man's mind, as always, two different schools of thought began to act. There is one positive approach, which is used and believed that the human mind is a computer character and therefore can be played in machines. Another school of thought took a negative stance and suggested that human intelligence cannot be replicated, and the role of computers is limited to assistance and testing tasks performed by people with the intellect.

These types of differences lead to confusion and sometimes cause problems largely; however, people must always welcome this type of discussion.

It is this difference of opinion, which caused most major things that need to happen on this earth. Here also, the difference of opinion brought another set of people in the arena, who invented in the middle of the path and AI, as a representation of the process of developing systems that reflect the human mind in their approach within a limited area. They are happy as long as computers operate their interest and can be used to solve a specific set of problems. It is not necessary for them to worry about the computer show the total exploration and development of the universal intellectual system .Hence, for majority of people,

"*AI is the branch of engineering employed for the creation of computers that possess some form of intelligence and can be used to solve real world problems and function within a limited domain.*"

It cannot be claimed whether this is the appropriate definition of AI. Different researchers have proposed different definitions of AI based of their own study and understanding of the subject. Some of these are given below:

"*AI is the automation of activities that we associate with human thinking, activities such as decision making, problem solving, learning....*"-(Bellman, 1978)

"*AI is concerned with designing intelligent computer systems which exhibit the characteristics we associate with intelligence in human behavior*" –(Barr and Feignenbaum, 1981)

"*AI is the exciting new effort to make computers think...machines with minds, in the full and literal sense*"-(Haugeland,1985)

"*AI is the study of mental faculties through the use of computational models*"-(Charniak and McDermott,1985)

"*AI is the art of creating machines that perform functions that require intelligence when performed by people*"-(Kurzwell,1990)

"*AI is the study of how to make computers do things at which, at the moment, people are better*"-(Rich and Knight,1991)

"*AI is the study of the combinations that make it possible to perceive, reason and act.*"-(Winston,1992)

"*Computational intelligence is the study of the design of intelligent agents*"-(Poole et al.,1998)

"*AI...is concerned with intelligent behavior in artifacts*" –(Nilsson,1998)

3.2 ELIZA Speaks.

One of the most impressive early successes in the field of artificial intelligence was ELIZA. ELIZA was computer software invented by Joseph Weizenbaum (1966), professor of computer science at MIT. Invented in 1966, the ELIZA was able to respond to the information that was entered into the computer.

ELIZA responses were simple. They were similar to those of a doctor or a psychologist to use when gathering information. For example, if someone typed in "I have a headache," ELIZA answer: *"Why have a headache?*" ELIZA responses were always questions, but it was for other computer scientists that they could create more complex forms of artificial intelligence.

3.3 Turing test.

The Turing test, proposed by Alan Turing (1950), was designed to ensure a satisfactory working definition of intelligence. The computer passes the test if the person is the investigator, after posing some written questions, do not say, and do written responses from the person or computer. In this test, it discussed the details of the test, and whether the computer to do it would be reasonable, if it is passed. Now it is clear that programming is so strictly apply the test gives you a lot of work. The computer must have the following characteristics:

natural language processing to enable it to communicate successfully in English;

knowledge representation to store what it knows or hears;

automated reasoning to use the stored information to answer questions and to draw new conclusions;

machine learning to adapt to new circumstances and to detect and extrapolate patterns.

Turing's test deliberately avoided direct physical interaction between the interrogator and the computer, because *physical* simulation of a person is unnecessary for intelligence. However,

3.4 Voice command process.

Generally, there are two main aspects of voice command processing:

– Speech recognition (phoneme or word recognition).

– Language processing (parsing, semantic analysis, pragmatic analysis).

Each of these research domains, has its own problems, solved and unsolved problems, so different research groups often treat them separately. Sometimes they are treated as separate, independent steps in voice command processing. As it will be shown later in this article , this approach is not justified , and - except for a very simple, functional bounded solutions cannot be used to develop real systems, voice control , which can be implemented and practically tested .

As far as speech recognition algorithms are concerned, it should be clearly stated that in modern engines SR- not as reliable as the human ear. Therefore, different methods are used in order to improve their reliability. One of them is user training voice profile (in the case of speaker-dependent motors SR). It is doubtful, however, whether speaker-dependent solutions to correct for industrial applications.

Some researchers have attempted to compare the different speech modules for their effectiveness. For example, some of them provide test results, showing that Microsoft SAPI is one of the best products, which is characterized by very low recognition of the time and very high recognition accuracy. Not surprisingly, the Microsoft SAPI has been used in voice control systems by many researchers. Even if a very high recognition accuracy reported (approximately 98%), there is still a problem that any misunderstanding can be important (sometimes even a catastrophic effect on the functioning of the production system.

Moreover, in my experience, high recognition accuracy rate according to some authors, it can be achieved only under certain specific circumstances (e.g. the use of high quality microphones and an absolute lack of noise). Moreover, note that the level of recognition of a particular speech engine often depends on the volume of the engine installation. Example, Microsoft SAPI allows the user freely adjust the pronunciation sensitivity and recognition. In depending on the current parameters (number of possible values theoretically infinite) recognition accuracy may be vary.
Another set of problems in the area of speech recognition (although partially overlapping with the domain of language processing) arises due to the spontaneous speech effects. Some of those effects are as follows:

– Pauses in speaking due to the speaker's hesitation: sometimes they may be mixed up with the end of the utterance.

– Word repetitions and inserted words with no importance for sentence meaning.

– Omitted (or spoken unclearly) words significant for sentence meaning.

From a linguistic point of view of treatment, the most important issue is the selection of possible commands formulas that are characteristic of natural language. This variety has several sources. One of them is the presence of synonyms, i.e. different words having the same meaning. Another variety of syntactic structures. It is not so obvious in English, but is a big problem in inflected languages. Solutions presented in this document is universal (independent of language) and considers this issue as well. Figure 1 depicted the process of voice command.

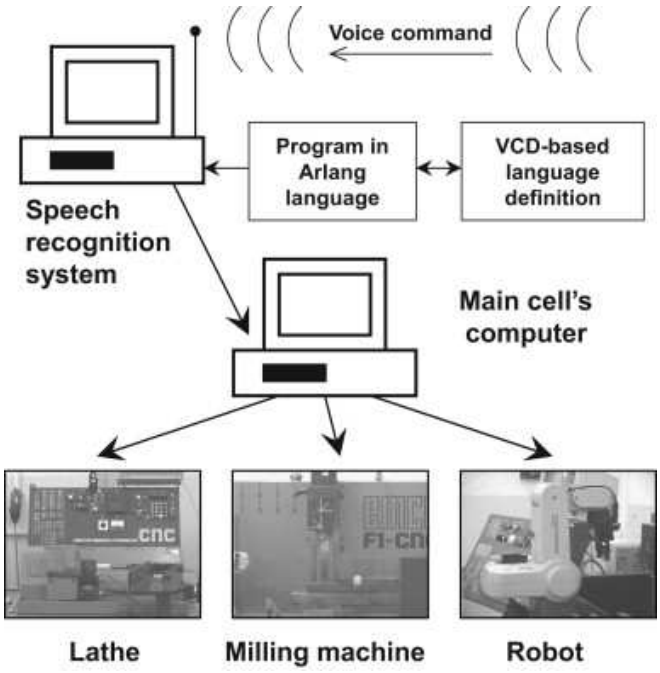

Figure 1:Voice command process

(Source: http://www.sciencedirect.com/science/article/pii/S0736584511001189)

3.4.1 Specific features of industrial voice control systems.

As mentioned earlier, existing speech recognition and natural language processing system, having lost control of any industrial equipment. Voice control of applications that are used to control industrial devices are but a minority. The reason, obviously, is not lack of interest, rather than the belief that the voice - it is not a suitable way for man-machine interaction and communication in industrial environments. For example, Kulyukin (2010) provides practical and scientific arguments for voice interaction between man and machine. Pires (2009) shows the importance of voice in industrial robotic cells in the future scenario where humans and robots to communicate securely workspace and actively cooperate. Consequently, the problem is likely in the absence of reliable solutions that could be applied in actual production conditions. To solve this problem, it is necessary to ask about the specific opportunities of industrial systems voice control. The answer to this question will

help us decide what elements common Smoking industrial solutions can be adapted and what needs to be redeveloped.

It could be noticed that some requirements that must be fulfilled by industrial voice control systems are more restrictive than those related to the non-industrial solutions are. The most important ones are the following:

– Correct recognition of all the words of the voice command is particularly important. One falsely recognized word could be a reason for a wrong (and even dangerous) action performed by machine.

– The language used in engineering is rich in expressions representing numerical values. Wrong recognition of words representing numbers will be sure to result in wrong actions.

– Reaction to the voice command must be usually instant.

On the other hand, some requirements, significant in other domains, are not so restrictive in industrial voice control applications. Here are some examples:

– As the users of those applications will be qualified machine operators, they can be expected to adapt themselves to some restrictions regarding the voice command structure. Therefore, a quasi-natural language could be used instead of the natural language. Nevertheless the language processing should take into consideration both lexical and syntactic diversity of voice commands.

– Unlike the users of assistive robots (who are often non-professionals or even people with disabilities), the machine operators can be expected to keep some discipline in speaking e.g. to avoid repetitions and insertions.

– Although the number of possible command formulations may be extremely high, the number of corresponding qualitatively different actions to be performed by machine is usually quite low. As it will be shown later, this fact influences the way the voice commands grammar can be defined.

3.4.2 Basic concept of industrially oriented voice control system.

The analysis in the previous sections shows that the voice control based on NLP does not seem appropriate solutions in industrial oriented applications. A similar conclusion with respect to Kulyukin(2010) subsidiary robots domain . He argues that the speech recognition algorithms are not robust enough to ensure error-free control of a robot using voice commands in natural language. Sublanguage acquisition is proposed as an alternative. The experiment was conducted in order to prove that a subset of people to learn languages quickly. While it is desirable to use a natural language voice control applications, the appropriateness of NLP solutions limited by many factors. Therefore, systems based on formal grammar appears more promising. A study Stiefelhagen(2009) et al. shows that the sentence grammatical error rate better than for n- gram -based speech recognition. However, as stated by Lauria(2010), mainly for the grammar -based approaches to the generation grammar with a wide enough coverage for the domain . Therefore, the author of this work focused on the question of how to take into account the great variety of possible formulations, using the command CFG. For this purpose, a special format voice command description VCD was created. It will be presented in the next section.

The industrially oriented voice control system developed by the author is based on the following assumptions:

– It should be easy to adapt it to any industrial robotized cells. Unlike most of other solutions described in the literature and it is not application oriented but it provides a generic tool for creating voice control applications.

– CFG-based language model should be used but diversity of possible command formulations must be taken into account.

– Requirements specific for industrially oriented systems (listed in one of the preceding sections) must be fulfilled.

– Spontaneous speech effects should be taken into consideration to some reasonable extent.

The system consists of three main modules:

– Speech recognition module (including vocabulary).

– Syntactic and semantic analysis module.

– Execution module (responsible for controlling machines and devices).

Because the system must be easy adaptable to any robotized cell, it fulfills the following conditions:

– Speech recognition module is universal (completely independent of particular application).

– Syntactic and semantic analysis module has generic character. There are fixed algorithms for parsing and for meaning extraction hence they do not have to be developed for each new application. However, sublanguages used for expressing commands in individual applications can be defined using textual VCD format. The same VCD format is used for command meaning description. This is the main difference between this system and the other ones known from the literature.

– Execution module is (of course) application oriented and must be developed for each application. However, a novel arithmetically oriented semantic analysis algorithm (that will be described later in this paper) simplifies the communication between (fixed) semantic analysis module and (application oriented) execution module.

Adaptability decision was practically tested in laboratory studies at the Warsaw University of Technology. Students without professional experience in voice control were able to develop simple applications very quickly.

Requirements specific to industrial oriented voice control systems are implemented. Use regular grammar (CFG) limits the number of statements that can be expected, therefore, wrong hypothesis can be rejected, and the recognition reliability, increased significantly. Numeric expressions that are part of the voice commands can be easily identified in the VCD. Examples thereof will be discussed later in this paper. Use regular grammar has another important advantage. Although it is possible recaps much, they are uniquely defined. Therefore, in most cases people can instantly recognize the end of a voice command without having to apply any "performance of speech." Thus, the instant reaction to the team can be achieved even if there is a pause (e.g. due to vibrations) in the operator expression. This is very important in industrial oriented voice control systems.

Some other spontaneous speech effects were also taken into account. Using the words or phrases (i.e. words or phrases irrelevant for sentencing value) is acceptable if those words or phrases included in the VCD-description as conditional elements. Ignoring unrecognized words or phrases as possible, but not recommended.

There is an opportunity to break the team mistakenly delivered and talk again. Appropriate mechanism would be presented in the next section.

3.5 The role of speech technology in mobile applications.

3.5.1 Just Say What You Want

Simple user guide "*just tell me what you want*" (SWYW). This, of course, ideal, and it is difficult to achieve, as in full generality and limitations of speech recognition integration with mobile operating systems and applications, as they are now.

In a voice user interface, will be functional categories, such as moving the application on phone; connect to the service based on a set of contact numbers, conducting web searches, dictating a text message, etc. The user can quickly learn that words like "*search* "," *dial* "or" *dictate a message* " will make the results more reliable and in the interpretation of tasks , at least a general message context restricted domains , it can be processed. He can do that as well say «*I do not understand*» if it cannot be classified a request to one of these domains. This feedback will help the user to find out what works. It is assumed that the user can quickly find out the scope for the team when it is necessary, until a certain way, in the context was provided was flexible and intuitive.

3.5.2 Talking to the Text box.

One model of voice user interfaces for mobile phones today is the ability to dictate in any text field, or, in some cases, in any part of the application that allows entering text. Some voice applications also have their own box that appears when start the application, perhaps by clicking on a button on the phone, or icon on the screen. Voice application, user has the option of interpretation voice commands run the application, the appropriate commands, and does not require manual navigation for this application.

Another model available is dictating into a "*clipboard*" application that is part of the phone's operating system. The contents of the clipboard can be pasted into most applications.

One deficiency of these approaches as they have been implemented to date is the lack of interactivity. Once one says something in the text box, some action is generally accepted that one drop of speech interface, e.g. displaying a list of search results. No dialog box. Dialogue is a powerful tool for disambiguation. For example, in web search, there are often a number of ways to interpret a request, and a long list of many non- option is the most difficult to view on user's mobile phone. Ideally, these interfaces will be developed to enable more dialogue when it might be useful.

3.5.3 Dictation.

While team-a request for a mobile phone can be implicitly limited, dictation of email, text messages or voice memos essentially unlimited. Dictation of free-form text version supported BZ number of companies, usually with a small client application on your mobile device and speech recognition within the network. This approach uses the data channel, and the speech quality that can be delivered via the data connection, usually better than the voice channel.

A key difference in dictation and voice requests is that the dictation is intended to be read and understood by a human, not a computer. The composer can edit the message before sending it. The computer or mobile device must in contrast, understand a voice command. Thus, the task for dictation different. It is harder in some ways and easier in others than a "*say what you want*" user interface.

The mobile phone being a personal device eases the dictation task. Most dictation systems tune both to the vocabulary usage and to voice of the user. One dictation application for mobile phones downloads and incorporates contact names and thus be accurate in transcribing proper names that are in the contact database.

3.6 The First Mobile Voice Applications.

3.6.1 Voice Dialing and Voice Commands on Phones.

One of the first mobile voice applications that appeared in the 1990s was the voice dialing , which allows users to press a button and say the number or name of the caller , so that the user can place the phone without looking at the keyboard , and trying to find the numbers.

Processing power and memory capacity continued to increase on mobile devices. Device manufacturers will be added in the near future more complex phonetic speech recognition device. Phonetic speech recognition using acoustic speech recognition models trained on a variety of styles and narration, and recognized phonemes and not templates word, and had the following advantages:

- No user enrollment or training was required
- New words and contact names could be added dynamically. If a new contact name is added to the contact list, then the voice dialer using standard phonetic pronunciation rules could recognize it.
- The application could recognize flexible manners of speaking. For example, a user could say "*Call John on his mobile*", or "*Dial John Smith on his cell phone*". If the application was correctly programmed, it could handle a great deal of flexibility.

Some voice command applications could also be programmed to recognize a long list of commands, beyond just dialing. In fact, some phones today can recognize 50-100 voice commands to control the device. Popular hands-free commands include:

- "turn Bluetooth on/off"
- "send text to <contact name>"
- "check battery"
- "check signal"
- "go to camera"
- and more

3.6.2The Advent of the Hands-free Experience on the Phone.

Voice dialing and other voice commands needed to work in situations where the user hands and eyes were not completely free, and it was therefore important that these applications could be subject to a minimum of attention interface.

Artists in speech recognition systems on the device must consider the amount of press and hold, want to use speech recognition. The easiest and safest interfaces required only a simple push of a button, as described in the following sequence:

- User pushes a button and quickly releases it to activate voice commands.
- The voice command application prompts the user via an audio cue to begin speaking.
- The user says, ”*Call Jim on his mobile*”
- The voice command system automatically detects when the user has finished speaking and begins the dialing process.
- If any disambiguation is required(for instance, if there are multiple entries for “Jim”), the voice command system resumes listening without requiring another button push from the user.

In addition, it is important that these words dialers provide audio signals to the user when they were not looking at the device. Audio cues and high quality text-to-speech were included in some applications, providing audio confirmation name / number to be dialed, and to disambiguate if there are multiple matches. For example:

User: “*Call Jim on his mobile phone*”

System: “*Multiple matches found...Jim Ardman...Jim Smith...Jim Workman”*

User: “*Jim Smith*”

System: “*Calling Jim Smith's Mobile phone*”

Text-to-speech must be used in this example to playback names from the contact list. If high quality text-to-speech is embedded on the device, then it can be used to enhance the minimal attention interface by performing services such as:

- Announcement of incoming caller ID number or name
- Announcement and reading of incoming text messages.
- Announcement and reading of incoming emails.
- Reading menus aloud.

Over the past few years, the manufacturers of the devices were deployed applications phonetic speech recognition system and high quality text-to-speech. One example is the

Nuance Vsuite product, which can support dozens of languages, and contact lists with thousands of names. These applications work best if they are integrated as fully integrated capabilities of the device to ensure maximum comfort.

3.6.3 Assistive Applications on Mobile Devices.

Speech technologies have been used on mobile devices to enable and enhance service for blind and visually impaired users, as well as those in the disabled community.

Common applications included:

– Voice dialing with audio confirmation

– Screen reading

– Caller ID announcements

– Reading incoming text messages and email

Helper programs necessary to address the needs of the user community carefully. For example, Nuance Communications TALKS screen reader for mobile devices functions, adjust the volume and speed of speech text-to-speech, and also included integration with external Braille input / output devices.

3.6.3 Emerging Mobile Voice Applications.

In the past few years, a number of new applications for mobile devices having that use voice-based network technologies. In some cases, these applications were available for download, or high-end smartphones such as the iPhone, Blackberry, Android, Symbian and Windows Mobile devices. In other cases, they pre-loaded onto the mobile device in a car or platform.

3.6.4 Voice Navigation and Mapping.

Application providers that make navigation and mapping technologies were among the first to introduce advanced technologies in speech processing their applications. Speech recognition technology is used for input / output easier when on the move, or by using small dimensions of the keyboard or touch screen keyboard.

These applications can be enhanced by:

– Entry of destination address by voice

– Entry of landmark or point of interest by voice

– Lookup business names or other content criteria (e.g., “Dave Matthews concert”)

– Playback of specific turn-by-turn directions using text-to-speech

Implementing speech enabled navigation can be challenging, especially for multilingual systems. Department of speech recognition technology itself is not enough. Technology should be trained on the "long tail" of addresses and places that people will need directions. In addition, it is important to support applications natural language interfaces, as users will have a low tolerance for the next few steps to enter the city, state, and say the names of companies or destinations extremely restricted mode.

3.6.5 Message and Document Dictation.

The emergence of text messaging and e-mail as a popular mobile applications fast, partly due to the presence of a full QWERTY keyboard mobile devices. However, small keyboards and are difficult to use for many users, touchscreens difficult to use for typing, and it is impossible and unsafe in on-the-go situations.

Over the years, the text was dictated by the successful application in the desktop and in the laptop world with software like Dragon Naturally Speaking, which is trusted and used by millions. Network based computational capabilities allowing currently convert speech to text dictation to mobile devices. Nuance recently released version of Dragon Dictation for the iPhone, which provides a simple user interface to dictate text for e-mail, text messaging, social networking applications, and any application that requires text input.

Dictation technology will work best when integrated into the applications that use dictation, such as email and messaging clients. On some mobile operating systems, such as Symbian and Android, it is possible to include speech as a universal input method that is active in any application that allows text entry. On feature phones and other operating systems, it may only be possible to include speech dictation by modifying the applications that need to use dictation to interact directly with the recognizer.

There are several important ingredients for success of speech dictation in mobile applications:

– The speech-to-text technology must be mature and robust for the language, which is being dictated. It can take years of real-world use from a variety of human voices to make this technology robust

– The user interface must be clear about when and how to activate speech recognition

– Ideally, the speech recognition technology can learn from the user's voice, common names they use, common terms used in their email and messages…this can require the user to give permission to upload some of this data to the network

– The user must have some way to correct mistakes; ideally, this will be a "*smart*" correction interface that gives the user alternate word/phrase choices so they do not need to retype.

3.6.5 Voice Search.

Similar to voice dictation, voice search allows the user to perform search queries using their voice. These could be:

– General search queries fed into a search engine such as Google, Bing, or Yahoo

– Domain specific queries, such as searches for music or video content, product catalogs, medical conditions and drugs, etc.

For voice search works well, speech technology, should be trained in common terms used in search queries. Total voice search engine should know about celebrity names, top news topics, politics and other medical voice search engine, should be trained in medical terminology and drug names.

Voice search has been built in many popular search engines. However, it may become more interesting as the applications emerge, which can determine the type of search and the user's intent and start a search relevant content source.

3.7 Speech and the Future of Mobile Applications.

Enterprise Applications and Customer Service.

Businesses such as banks, mobile operators and retailers have begun to invest in mobile applications. Rapid introduction of smart phones such as iPhone, Blackberry, and Android-based phones, provided a set of development platforms downloadable applications, which can lead to a wide segment of the customer base.

Speech recognition offers many advantages for maintenance applications today in a phone with voice applications. These advantages can be extended to mobile customer service applications, so that callers can speak for mobile applications to get information about the product, account, or technical support. It can remove restrictions from mobile usability interface and enable enterprises to build more complex applications that deliver the best self-service capabilities.

Potential examples of speech usage would be:

– Using an open-ended "*How can I help you?*" text box at the beginning of the application that would enable the user to type or speak their question and then launch an appropriate mobile applet (a small application that performs limited tasks) that would provide service…instead of forcing the user to navigate through visual menus.

– Adding a product search box to a mobile application, and so the user could say the name or description of the product for which they need service.

– Speaking locations for store/branch locators.

– Speaking lengthy account numbers or product codes for service activation

– Dictating text into forms for applications (e.g., a mobile loan refinancing application).

Companies may find valuable use for mobile workforce applications, such as:

– Dictating notes into CRM applications

– Dictating notes into work order management

– Dictating into mobile order processing applications

3.7. «Your Word is my Command": Google Search by Voice: A Case Study.

3.7.1Google Maps for Mobile (GMM).

Traditional directory assistance applications are restricted to one modality using voice as input and output. With the advent of smartphones large screens "and the connection data, we could move to a multi-modal user interface through speech or text as an input modality, and maps with super-imposed business proposals as output modality.

Figure 2:Google Maps for Mobile, with voice interface.

(Source:
http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/36340.pdf)

In March 2008, we released our first multi- modal applications for word Google Maps for Mobile (GMM). Fig. 2 shows a multi- modal interface to support the directory that we have built on top of TMG.

Multi - modal experience has certain advantages over the IVR (voice) system. First, the output may be a visual modality, and as mentioned, which allows for much richer flow of information. Google Maps for MO bile can see the location of business and other related information directly on the map. Contact information, address, and any other meta information about the company (e.g., ratings) can be easily and etc. A second major advantage for the time it takes the user to search "and digest the information. Because multimodal search experiments hence, the total time taken is much less, than that after a single, I / O say modality IVR system. Finally, the cognitive load on the user significantly reduced - ephemeral speech imposes very serious cognitional requirements of the user when information is transmitted or prolonged complex. These advantages can significantly improve the quality and quality of interaction information can be provided as compared with traditional systems IVR.

3.7.2 Google Search by Voice

Mobile Web Search is a rapidly growing area of interest. Internet-enabled smartphones account for a growing share of mobile devices sold worldwide, and most models experience of viewing web pages that rivals desktop in display quality. People are increasingly turning to their mobile devices when involved in a web search, e driving orts to increase the usability of web search on these devices.

Although mobile devices improved usability, typing search queries can still be cumbersome, prone to errors and even dangerous in some Frequency scenarios.

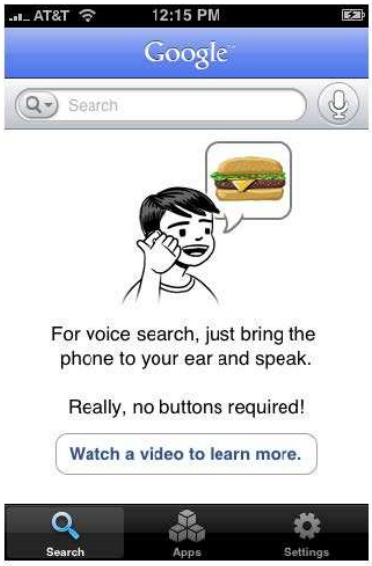

Figure 3:Google search by Voice for iPhone.

(Source: http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/36340.pdf)

In November 2008, we introduced the Google Mobile App (GMA) for iPhone (Figure 3) that enabled voice search. GMA Voice Search advanced paradigm multimodal voice search, search business places on the map to find all the World Wide Web. In the next few sections, we will discuss the technology behind these e orts and some lessons we learned by analyzing the data of our users.

4. Principle of mobile application.

4.1Voice Recognition and Commands – Background.

Speech recognition (SR) is the translation of spoken words into text. It is also known as automatic speech recognition (ASR), computer speech recognition, speech to text (STT)

Speech recognition applications include voice user interfaces such as voice dialing, call routing (hands free), Demotic device control, search (search podcasts, in particular where it was said), data entry (transcription), preparation of structured documents (radiology report) , speech, word processing (word processors or emails) and aircraft (Direct Voice Input)

Speech recognition system over 10 years ago, are also faced with a choice between discrete and continuous speech. It is much easier for the program to understand the words when we say them individually, with a distinct pause between them. However, most users prefer to speak of ordinary conversational speed. Most modern systems capable of understanding continuous speech.

This report examines the issues of patent activity around voice control mobile devices, and contains the main arguments and NPE working in this area.

4.2 Speech Recognition Technology Categorization.

The records were further classified around the following two key areas:

By Algorithms: Hidden Markov models, Dynamic time warping, Neural networks

By Application Area: Defense, Healthcare, Navigation, Telematics, Home Automation, Games, Hands-Free Communication, Multimodal Interaction, Speech-To-Text Reporter, Robotics, Pronunciation, and Transcription.

4.3 Key Companies

• The chart shows top 15 companies (Figure 4) for voice recognition.

• IBM has the maximum number of filings followed by AT&T and MICROSOFT

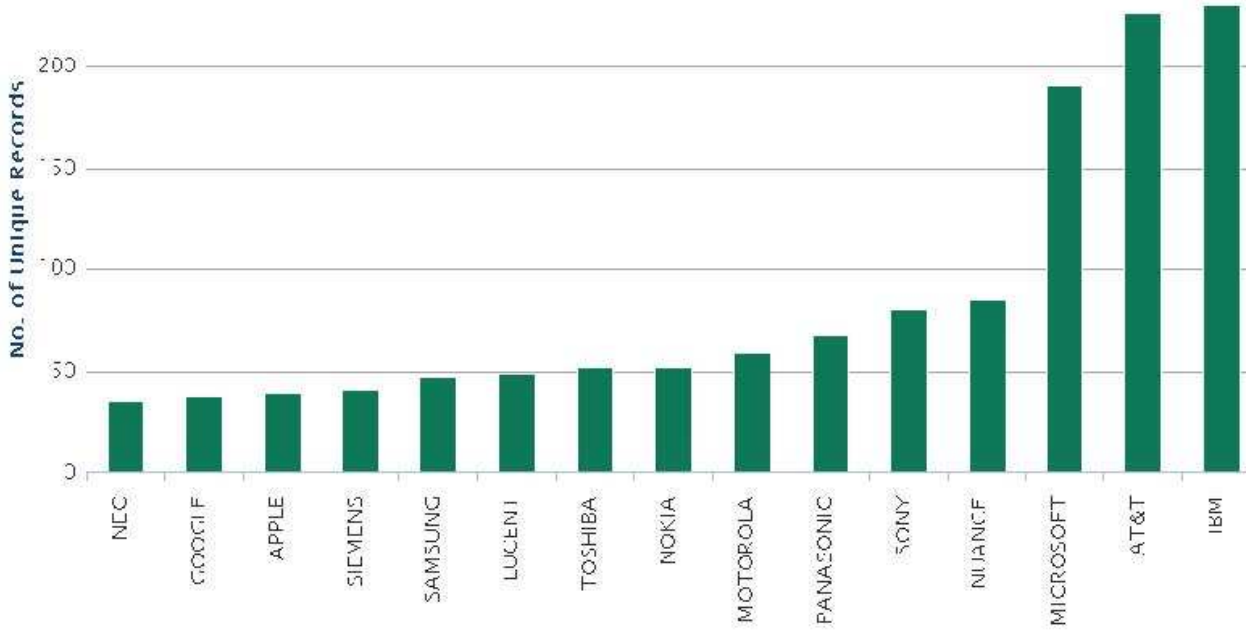

Figure 4: The chart of top 15 companies for voice recognition.

(Source: http://patseer.com/wp-content/uploads/2013/07/Voice-Recognition-in-Mobile-Devices.pdf)

Priority Country Map – Where is research being done?

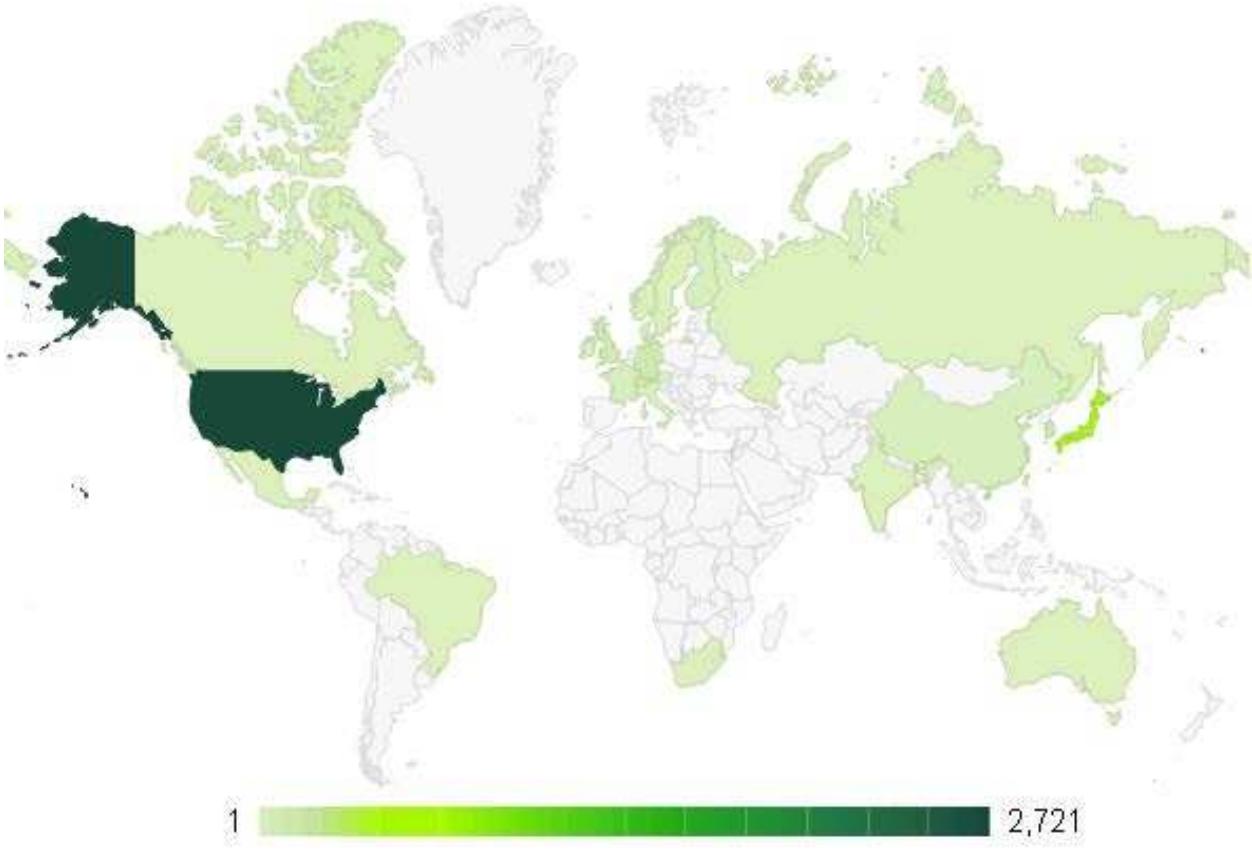

Figure 5: Geographical distribution.

(Source: http://patseer.com/wp-content/uploads/2013/07/Voice-Recognition-in-Mobile-Devices.pdf)

• The map shows the geographical distribution of filings for voice recognition (Figure 5)

• Patent protection is being sought mostly from US followed by JP

4.4 Using Voice and Speech Recognition with Mobile Devices.

Smartphones and other mobile devices are in the middle of the major innovations in technology to provide hands-free access to the functions and navigation, often called «*the voice commands, voice-enabled, voice actions or speech recognition.*" This technology has significant implications for people who have a disability, assistive technologies. Unless the user has a strong, clear voice these devices are becoming easier to use and provide increased access to the Internet, the use of mobile devices and the availability of communication. This

paper examines the use of the basic functions of a mobile device. When mobile devices with Wi-Fi many of these applications will increase the Internet using the device. Some applications or programs will require the device to be connected online to work. Others will work without an Internet connection.

4.5 Accessibility Feature, Mobile Device

Figure 6: Symbol for "tap to speak

(Source: http://www.atarizona.com/docs/General/VoiceSpeechRecognitionMobileDevices.pdf)

4.5.1 What is a "native" accessibility feature within a mobile device?

For Voice and Speech, born there must be some way to use voice commands and a way to dictate commands by using built-in microphone. Voice commands are used as navigation tools to manage the device functions. Speech recognition is an advanced feature that allows voice commands and voice input to navigate and perform operations such as using dictation and access to telephone calls voice dialing and text messaging. Speech recognition can use sophisticated software to enable dictation instead of touch-typing.

4.5.2 “Voice recognition” or “speech recognition”?

These two terms are often used interchangeably, but they really should not be. They have different values . Imagine that user is talking on the phone, listen for a few seconds and then say, *«Caroline, you can call me back? We have a bad connection. I can hardly hear you. »* What Voice Recognition, user could not hear it well enough to understand her words. Speech recognition systems, trying to understand the words spoken. Voice recognition can be used as a fingerprint for identification. No matter WHO said it?

4.5.3 Speech Recognition can be used to control a computer, navigate telephone menus, etc. What matters is WHAT was spoken.

The definition of "voice command".

Availability features including navigation devices made by using voice commands rather than pressing numbers or keys.

Using voice command means that there is speech recognition native to the device. In smartphones, the memory can be programmed to store and determine the names and spoke into the phone to call the number associated with them. Some smart phones and other mobile devices provide accessibility option is enabled so that the user can navigate to other settings, applications and menus with voice commands.

4.5.4 Voice Command for Mobile Devices.

Any mobile device running Google Android and Microsoft Windows Phone, IOS 5, or Blackberry OS provides voice command capabilities. In addition to native speech recognition software for every mobile phone operating system, the user can download a third-party voice applications teams from each of the operating system, applications from the store Apple App store, Google Play, Windows Phone Marketplace (initially Windows Marketplace for Mobile), or BlackBerry App World.

Other terms related:

Voice Control – Apple term for iPhones prior to the newest phones

Voice Actions – Android term for voice commands

Voice Input

Voice-enabled

Common native voice command/voice action or speech recognition tools in mobile devices:

VoiceOver – Apple products such as Mac computers and iDevices like iPod, iPad, and iPhone

Siri – iPhones only

Talkback – Android products like cell phones and tablets. Note that some speech recognition tools may require installing additional applications to make Talkback more fully-featured.

4.5.5 Accessories to improve or enable use of speech recognition.

All mobile devices that allow voice commands and voice recognition systems have a built-in microphone for this purpose. This may work well for some people, but those who have difficult positioning device for the microphone to pick up their voice will probably use an external microphone, headset with microphone or accessories that can be combined using a headphone and external microphone with Jack on hook together. Some typical external microphones in the figure 7, 8, 9 below:

Figure 7: Foam pad external mic

Figure 8: Thumbtack mic detail view

Figure 9: Capsule mic with cover,

(Source: http://www.atarizona.com/docs/General/VoiceSpeechRecognitionMobileDevices.pdf)

Most people, who need an external microphone for assistive technology will use a microphone/headset combination, or a plug-in microphone with a jack for the ear buds, all pictured below:

Figure 10: Typical computer headset with mic

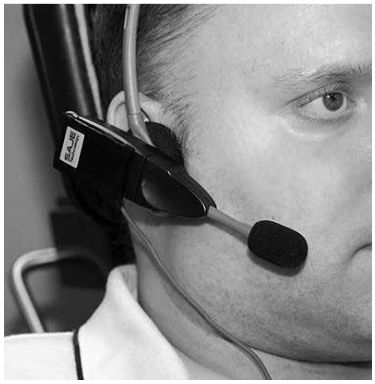

Figure 11: Bluetooth (wireless) headset with mic

Figure 12: Typical computer headset with mic

(Source: http://www.atarizona.com/docs/General/VoiceSpeechRecognitionMobileDevices.pdf)

4.5.6 Custom assistive technology accessories to improve use of speech recognition.

Ideal use of computer programs and applications for mobile speech recognition will use Bluetooth technology wherever possible. This build function of the system to enable the transmitter and receiver for this purpose, at least to enable the voice command for moving the device. Sometimes Bluetooth- it has just used to transmit sound to another device, and when the boom box speakers' project sounds much better than the device. Bluetooth is also used to connect remote devices with the base system. Good quality Bluetooth specifications allow the user to run more than one device that is "detectable" in the availability of the device settings, but the pair with the device quality via Bluetooth variable quality. Universal symbol for Bluetooth devices and programs listed below (Figure 13) as a trademark upper left:

Figure 13:Universal symbol for Bluetooth devices and programs.

(Source: http://www.atarizona.com/docs/General/VoiceSpeechRecognitionMobileDevices.pdf)

Both the base system of the computer or mobile device, AND the program must be compatible with the use of any accessories. It is important to verify specifications for all components that make assistive technology work for a user.

4.5.7 Free applications to improve speech recognition in a touch screen mobile device.

For what happens in the native mobile device, there is a free application for voice commands, and a full speech recognition. These sets are in the set, suitable for mobile phones android, iPhone and support for Wi-Fi tablets. Keep in mind that some applications only work on the device voice commands to navigate. Others may decide to issue more voice commands to work on websites. Some products can be designed for use only with performance, academic or business applications such as Microsoft Office Suite of programs. Many manufacturers and suppliers are working to create a speech recognition application for other products. Competition is very active at this time and in the future will bring many more options that can be even better for the use of assistive technologies. Speech recognition technology is one of the hottest trends right now, and for the future. Note that the operating system requirements for each device. For example, some applications work with

Android OS 2.0 or later, but others will not only work for devices with Android Jelly Bean OS version 4.0 or later. For iPhones, most work with the iPhone 4S and iPhone 5. The following free applications available information from the official websites of manufacturers supplying applications for smart phones and wi-fi -enabled mobile devices.

Examples of Android and iDevice applications:

Simple, free voice applications: Dragon Dictate and Dragon Search, Dragon Go! and Dragon Mobile Assistant. At this time, many of the Dragon applications are changing in availability and name. Any Dragon application made for mobile devices will work similar to each other (Figure 14).

Figure 14: Screenshots of a Dragon application with simple navigation.

(Source:
http://www.atarizona.com/docs/General/VoiceSpeechRecognitionMobileDevices.pdf)

A completely new idea for Dragon users:

Dragon Remote Microphone

Dragon Remote Microphone Check if there is an experienced user of Dragon with a computer program. This allows the user to operate the computer with another's smartphone or mobile device as External microphone to their computer.

App turns mobile phone or other mobile device into a wireless microphone, making it easier and more convenient for people to use their Dragon desktop. Version of Dragon Naturally Speaking, the computer must be current or recent. This application uses the Wi-Fi connection to PC when user run Dragon Software. Reviews are mixed at this time, with some enthusiasm, and others are disappointed. High reliability with 4S iPhone 5 January 2013.

This application can be used in Dragon NaturallySpeaking PC (Home Edition or higher for Version 11.5 or higher) and Dragon Dictate for Mac (version 2.5 and higher).

4.5.8 Browser-supported speech recognition, a new trend.

When working with a mobile browser to access the Internet, some of them will work in speech recognition device. Some Web sites and Internet access can be improved through the use of additional or substitute applications instead of built-in devices. Moreover, new technology has begun posting on the Internet large web sites with built-in speech recognition. Programming available online speech recognition tools can be done using special programming. To the public, Google is the first major companies to provide mobile voice search devices that can be used free downloadable application (Figure 15).

Figure 15: Google Voice.

(Source: http://www.atarizona.com/docs/General/VoiceSpeechRecognitionMobileDevices.pdf)

Voice Search is a feature of Google Search app for iPhone, BlackBerry, and Nokia S60 V3 phones. If owner has an Android phone, search for the "Voice Search" app in Google Play. A slightly different version, **Voice Search for Android** supports Voice Actions on Android 2.2 (Froyo) and above. The Android operating system also enables to speak instead of typing in any app where owner launch the keyboard. Voice Search is not available in all languages and all mobile phones.

5. Practical part

5.1 iPhone, Samsung, Windows Phone, Blackberry.

Software voice recognition is now a key selling point for smartphones. Apple gives a digital assistant called Siri, Samsung uses the power of speech via its functionality S-Voice, and Nokia, using the built in voice control within Windows Phone 8, receives phones actually listen this time.

IPHONE: SIRI

(iPhone 4S and later)

Siri is the name for voice control software created directly in the MOD, it is advertised as a digital assistant, able to schedule appointments, send text messages, add a calendar entry, and even do some searching Google.

All these features are available on the mobile, pressing and holding the home button, and after a few user seconds you will be presented with an interface Siri is simple, «what can I help you 'will appear along with the signal, so the user knows that he / she can start talking with his / her new best friend!

Siri could even be available from the lock screen in mobile, though it should be noted that some functions would be locked until the pin is not entered.

SAMSUNG: S - VOICE

(Galaxy S3 and later)

Samsung describes software as “intelligent personal assistant”, and from this point of view, as far as possible with S- calculated user’s voice. "Hello Galaxy" - a simple phrase that can be used to activate the phone and unlock the phone is ready to use, there is no pattern or PIN lock is on, of course. Unlike the iPhone and other speech recognition systems, for that matter, there is no need to press any key combination to activate the software, S- Voice will simply work when he hears someone say key phrases.

The voice command for apps is a powerful feature of S-Voice, it allows control over the following applications:

- Incoming calls
- Alarm

- Camera
- Music

Each of the apps have their own voice command control options. The screenshot of the handset's settings menu shows how comprehensive it is.

Author thinks that S-Voice is one of the more versatile voice recognition options out on the market today. Like Siri, it can be easily accessed via pressing down the home button twice, giving the user access to all manner of features. It can also be used for other functions that make it a more attractive option for those seeking voice controls, the ability to answer and reject calls and unlock handset by voice alone is a great step in the right direction for true voice control.

MICROSOFT/ WINDOWS PHONE VOICE CONTROL - SPEECH

(Windows Phone 7.1 and above)

Windows phones have their own version of voice recognition, which is available from the latest Windows 8 devices (as well as previous phones all the way to Windows Phone 7. Version from Microsoft called It and it may be accessed by pressing and holding the button (typically on Windows phones, usually the home button of the Windows logo) as Siri owner can say "Help" to have the phone to go into details as it can be used. Here is a breakdown of what it can do.

- Call any phone number
- Call someone from contact list
- Redial the last number
- send a text message
- Call voicemail
- Open an application
- Search the web

It can be used when the tube is blocked, this can be enabled in the settings, from here, the user can also include other accessibility options, such as talking caller ID, speed dial, turn forwarding on or off. One cool feature is the ability to read messages aloud - when the phone receives incoming text of man can say, "read" or "ignore” to hear the playback text.

The voice recognition system on Windows phones may not be as intricate as its competitors but it does allow the user to control core functions via voice, as well as the option to download applications such as VoiceMusic (free from the Market) to play music using voice commands.

BLACKBERRY 10 VOICE CONTROL

(BlackBerry OS 10 devices and above)

BlackBerry were not shy in including a voice recognition and control system in their new OS 10 operating system launched earlier this year. The imaginatively titled Voice Control app allows carrying out a number of basic tasks such as:

- Make a call
- Send an email, text message, or BBM message
- Search the Internet or BlackBerry device
- Book meetings and appointments
- Set reminders
- Change Facebook status or post a tweet
- Save notes and memos in BlackBerry Remember
- Set an alarm

The system is not digital assistant akin such Siri and S-Voice, but this is the basics of voice control very well indeed, and if he sends an SMS message or update account Facebook loud that after, this should be adequate. The user can access the system, press and hold the mute on physical QWERTY keyboard, if there is one, or by opening the Voice Control application from the applications menu.

5.2 Multiple-Criteria Decision Analysis

5.2.1 Ranking Method

Multiple-criteria decision-making or multiple-criteria decision analysis (MCDA) is a sub-discipline of operations research that explicitly considers multiple criteria in decision-making environments.

The challenge is to multi-criteria analysis of voice control in smartphones by the criteria mentioned below:

Technical characteristics:

- Language support
- Processing time of command "Call"
- Processing time of command "Dictate SMS"
- Processing time of command "Web Search"
- Processing time of command "Create task"

General criteria:

- Cost
- Population
- Availability

Quality criteria:

- Functionality
- Performance
- Reliability
- Stability
- Convenience
- Portability
- Scalability

Evaluation of three experts on the following criteria:

- Rating for technical characteristics
- Rating for quality

5.2.2 Algorithm

1. Rank each criteria category from 1 to 6, giving a 1 to the criterion that is most important, and a 6 to the least important.
2. Assign each criteria category a group weight by allocating 100 points among the 6 categories. The more important the criterion, the higher its weight. It's like grading a test, the sum of the scores for all test questions should add up to 100. Points given: 100
3. Author gives to each sub criterion its own weight. Weights can take on any value between zero and the maximum of the weight given to that group. For example, if assign a group weight of 15 to a criterion, the sub criteria in that group can range from 0 to 15. It can be assign all sub criteria in that group a 15, or some a 15, some a 10, one an 8, one a zero, and so on.
4. Calculate rating for each mobile phone

Criteria Categories and Subcriteria	Rank	Group Weight	Subcriteria Weight	Iphone	Samsung	Nokia	Blackberry	IPhone	Samsung	Nokia	Blackberry
Technical characteristics	**1**	**25**						**495**	**485**	**440**	**405**
Language support			25	5	4	3	3	125	100	75	75
Processing time of command "Call"			20	5	5	4	4	100	100	80	80
Processing time of command "Dictate sms"			20	5	5	5	4	100	100	100	80
Processing time of command "Web search"			22	5	5	5	5	110	110	110	110
Processing time of command "Create task"			15	4	5	5	4	60	75	75	60
General criteria	**3**	**20**						**165**	**185**	**160**	**120**
Cost			20	5	3	3	2	100	60	60	40
Population			15	3	5	4	2	45	75	60	30
Availability			10	2	5	4	5	20	50	40	50
Quality criteria	**6**	**25**						**585**	**585**	**510**	**360**
Functionality			25	5	5	4	3	125	125	100	75
Performance			20	5	5	5	2	100	100	100	40
Reliability			20	4	4	4	4	80	80	80	80
Stability			15	5	5	4	2	75	75	60	30
Convenience			15	5	5	4	3	75	75	60	45
Portability			10	3	5	5	3	30	50	50	30
Scalability			20	5	4	3	3	100	80	60	60
Expert #1	**3**	**10**						**100**	**90**	**90**	**70**
Rating for technical characteristics			10	5	5	4	4	50	50	40	40
Rating for quality			10	5	4	5	3	50	40	50	30
Expert #2	**4**	**10**						**85**	**76**	**77**	**66**
Rating for technical characteristics			10	4	4	5	3	40	40	50	30
Rating for quality			9	5	4	3	4	45	36	27	36
Expert #3	**5**	**10**						**72**	**63**	**72**	**63**
Rating for technical characteristics			9	4	3	4	4	36	27	36	36
Rating for quality			9	4	4	4	3	36	36	36	27

TOTAL SCORE	1502	1484	1349	1084
	IPhone	Samsung	Nokia	Blackberry

Figure 16: Multi-Criteria Analysis

(Source: author's research)

In this table, author analyze the Voice Control 4-x species Smartphones (iPhone, Samsung, Nokia, Blackberry) Appreciate every smartphone on the above criteria, and 3 experts give an estimate for each smartphone on their own criteria.

5.2.3 Analysis of the results by criteria

After evaluation of each smartphone, we get results. In this section, we analyze and compare the results obtained scores smartphones.

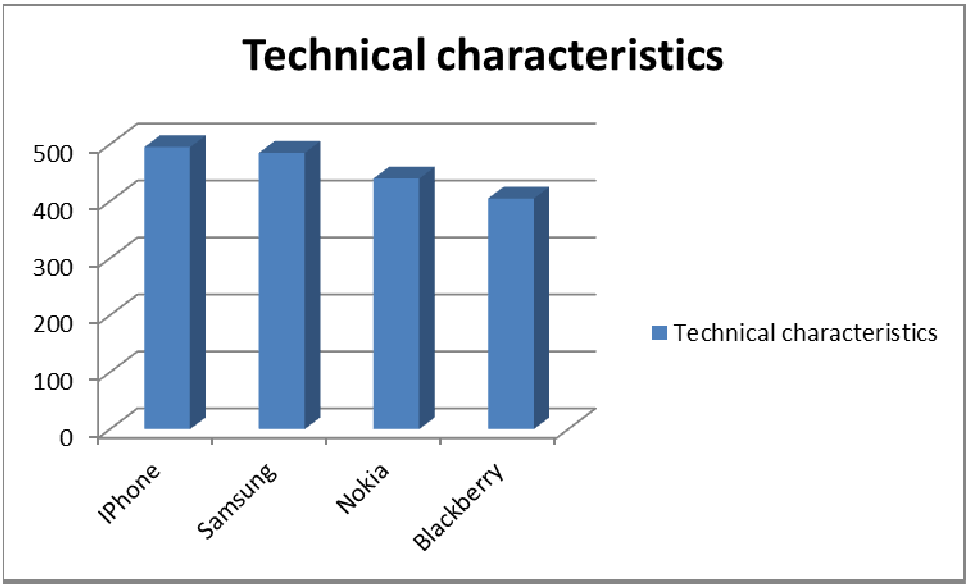

Figure 17: Graph derived summary scores on the technical characteristics.

(Source: author's research)

In Figure 17 it is visible that the specifications iPhone smartphones and Samsung smartphones are better than others.

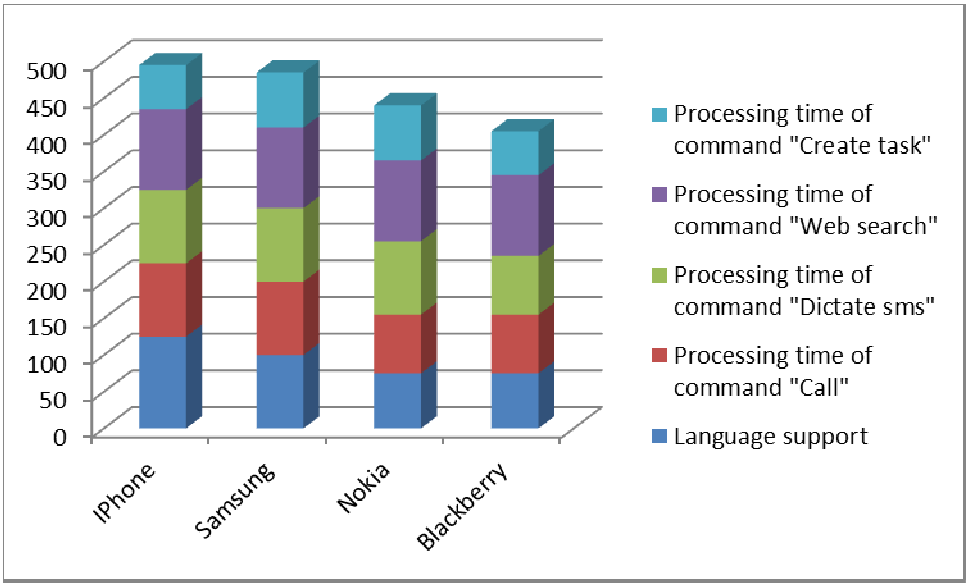

Figure 18: Schedule of grade points received by the specifications.

(Source: author's research)

In Figure 18, it is visible that the smartphone iPhone better language support other smartphones, i.e. it supports more languages than others.

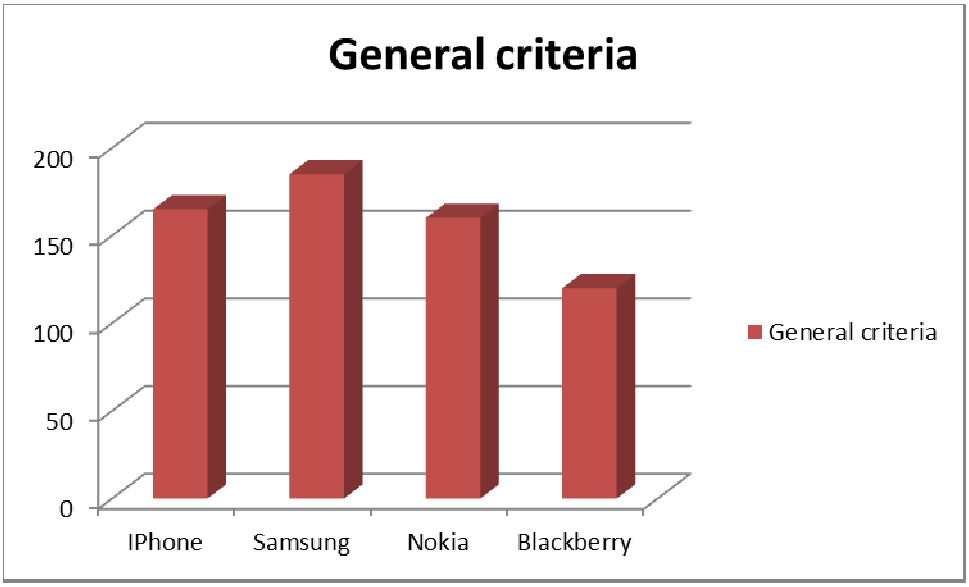

Figure 19: Graph derived summary scores on the general criteria.

(Source: author's research)

Figure 19 shows that by the general criteria leading smartphone Samsung. The last position owned the smartphone Blackberry.

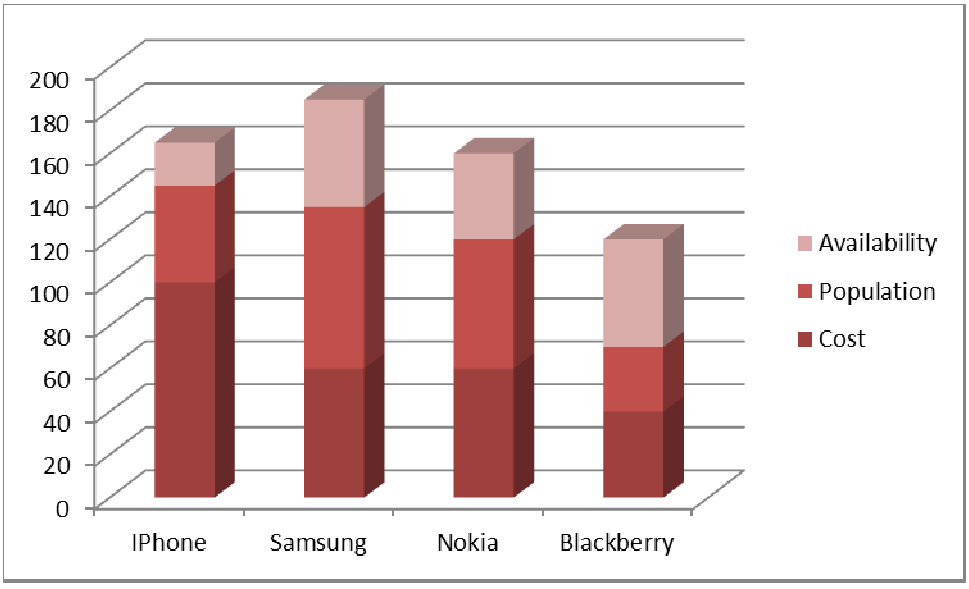

Figure 20: Schedule of grade points received by the general criteria.

(Source: author's research)

The graph (Figure 20) shows that the most popular smartphone is Samsung. By accessibility wins smartphone Blackberry.

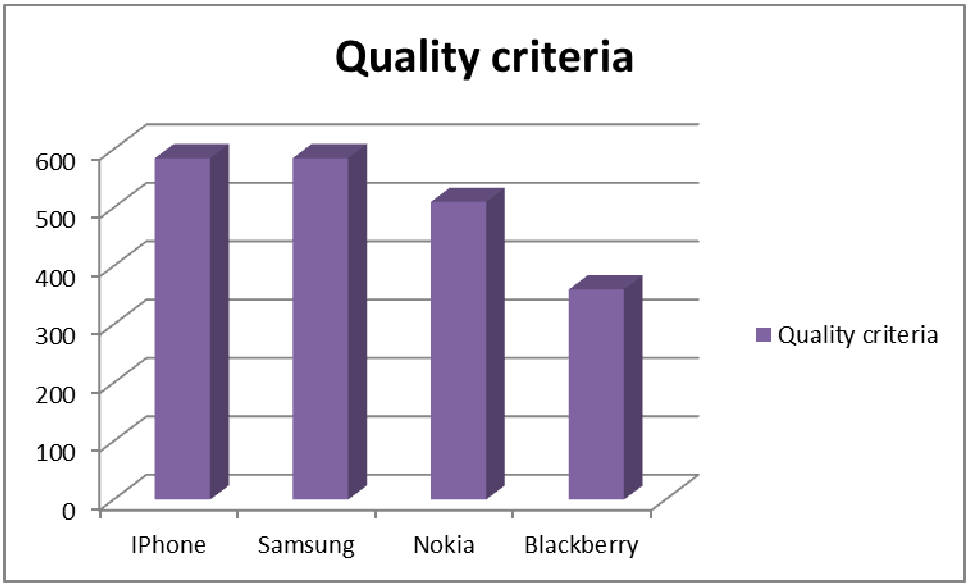

Figure 21: Graph derived summary scores on the criteria of quality

(Source: author's research)

This graph (Figure 21) shows that by the quality criteria lead Samsung and iPhone.

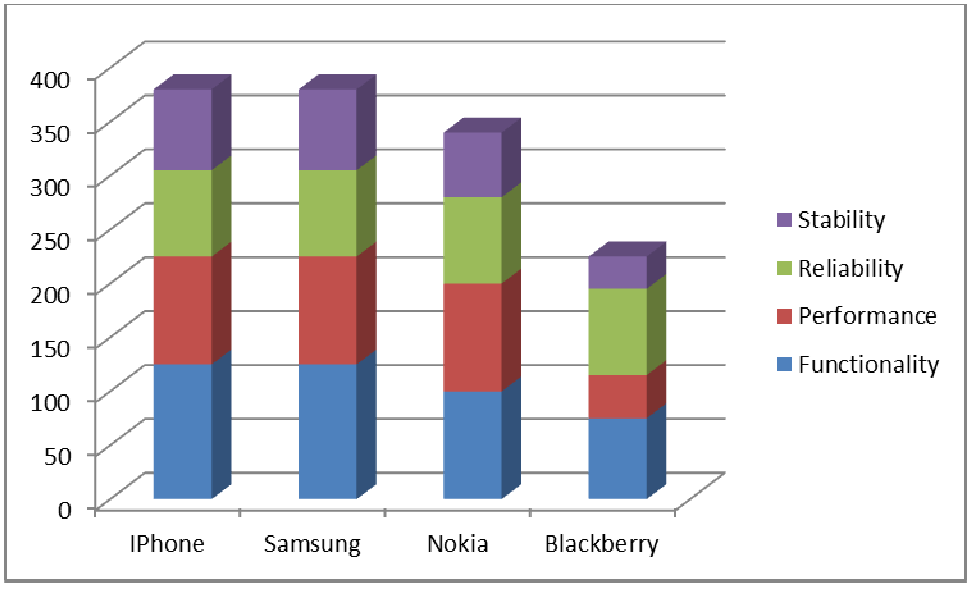

Figure 22: Schedule of grade points received by the quality criteria.

(Source: author's research)

This chart (Figure 22) shows a benefit in terms of functionality that Samsung and iPhone dominate.

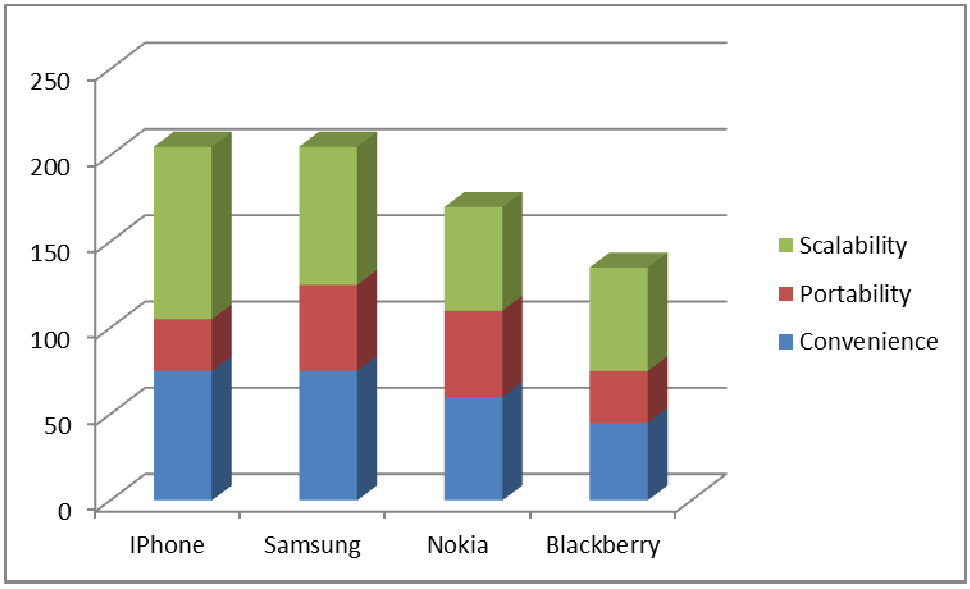

Figure 23: Schedule of grade points received by the quality criteria.

(Source: author's research)

In this graph (Figure 23) of scalability iPhone is leading.

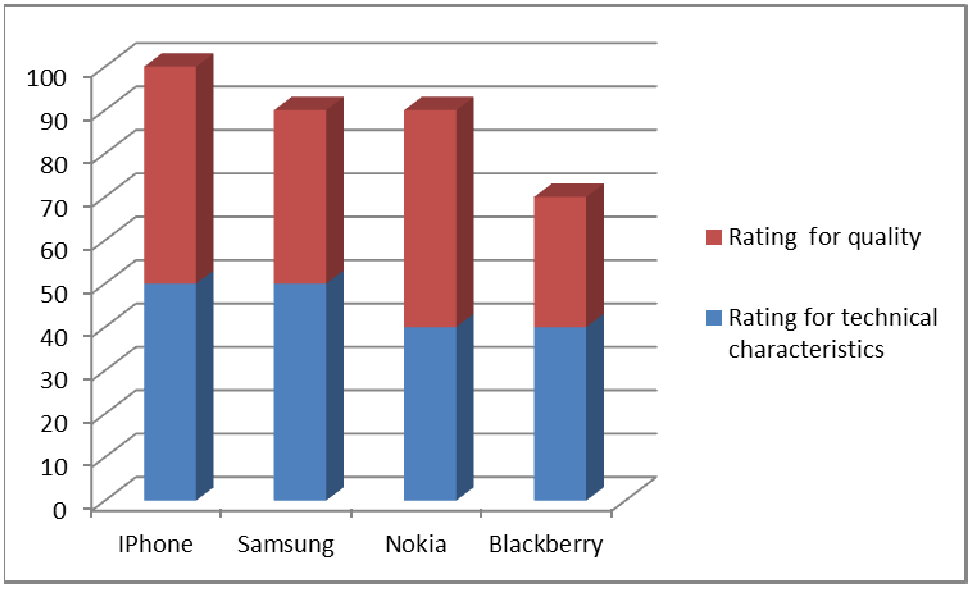

Figure 24: Evaluation of Expert №1.

(Source: author's research)

From Figure 24, it is possible to see expert's preference that gives to iPhone.

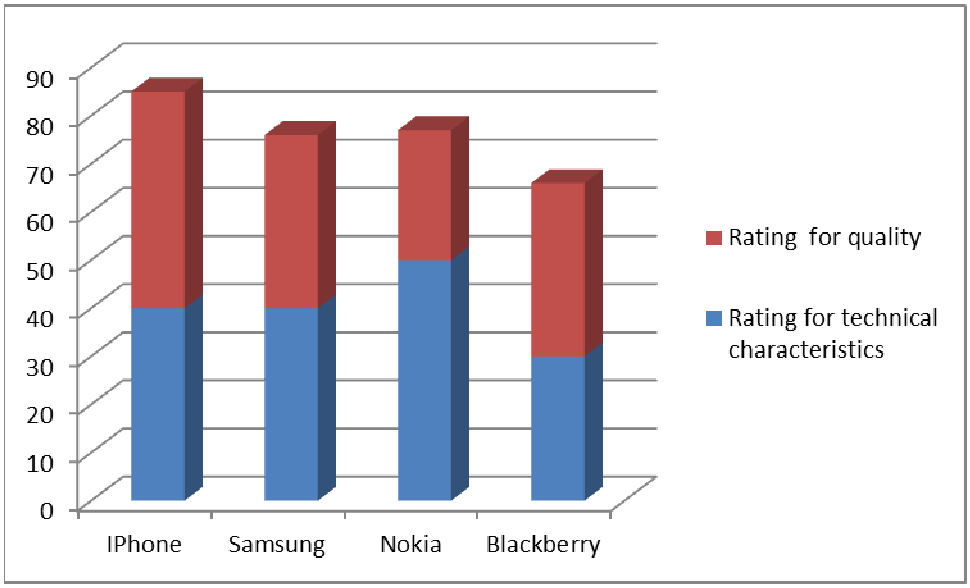

Figure 25: Evaluation of Expert №2.

(Source: author's research)

Expert№2 believes that Nokia is leading by the technical criteria (Figure 25).

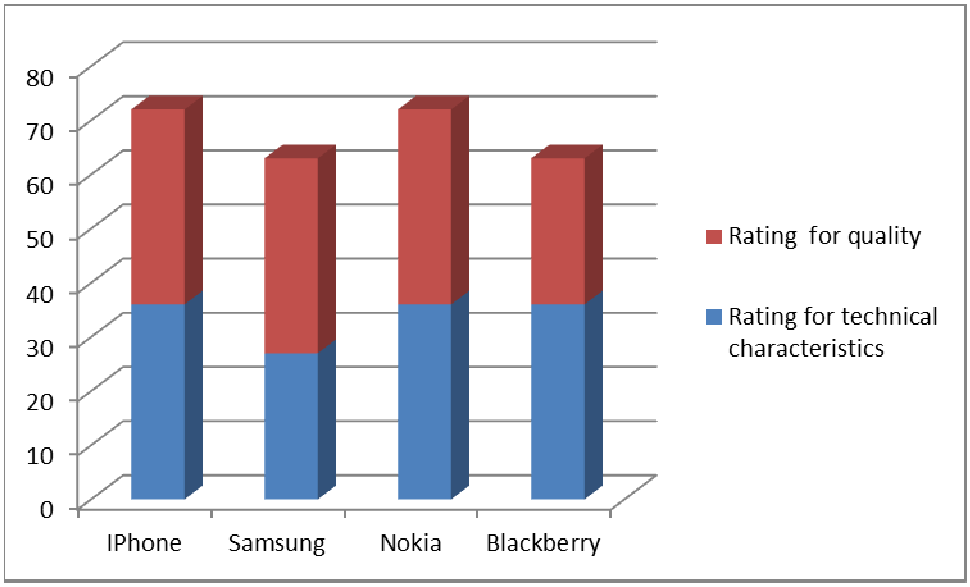

Figure 26: Evaluation of Expert №3.

(Source: author's research)

Expert №3 gives preference to iPhone and Nokia smartphones (Figure 26).

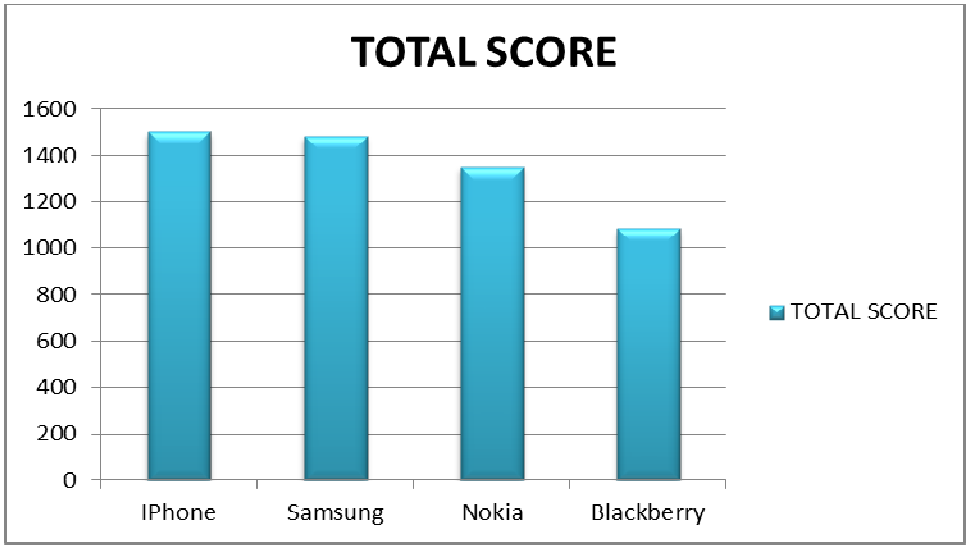

Figure 27: Combined scores for all criteria

(Source: author's research)

This graph (Figure 27) shows the total score for all criteria.

6. Discussion and results

Apple has added voice control device families IOS as a new sign firmware iPhone 4S,iPhone,IPad (3-rd and 4-th generation) ,iPad Mini , and iPod Touch (5th generation) all come with the new and more complex voice control called Siri , replacing voice control , which is still present on older devices. Siri user independent built-in speech recognition, which allows the user to issue voice commands. With the assistance Siri the user can perform the following voice commands : send text the message , about the weather , to set the reminder , find information , to schedule meetings, email , find a contact to set the clock , get directions , and track your stocks , set the timer , and ask participants to share examples of batch requests the voice tag. Other main S-Voice include the possibility of voice dialing , make a memo , schedule tasks , search contacts , make pictures , take calls , Internet searches , and tweeting , all with one voice and send e-mails and texts , adjust volume, and more .

Windows Phone operating system is the Microsoft mobile device. On Windows Phone 7.5, voice recognition is called speech. It the user-independent and can be used for : call someone from the contacts list , call any phone number , redial the last number , send a text message , call your voice mailbox, open the app , read meeting, request the status of your phone and search the web . In addition, it can also be used during a phone conversation, and the following actions are possible during a phone call: press a number, to turn the loudspeaker on or call someone that puts the current call on hold.

On the BlackBerry smartphone running BlackBerry OS 7 or later , it can be used to perform the following tasks : call the contact's name or number , to determine what phone number the user is currently using , check the level of signal , check the network , check the battery and it has the redo command. In addition, Blackberry OS allows the user to change the language for commands; the language of voice prompts and allows you to train the voice recognition.

7. Conclusions

Analysis of results was made by following:

- Based on the results of technical criteria, iPhone and Samsung are leading.
- Based on the results of general criteria, the leader is Samsung.
- Based on the results in terms of quality, iPhone and Samsung are leading.
- According to experts, leader is iPhone.
- By popularity, Samsung is the leader.
- By functionality, iPhone is the dominator.
- By availability, winner is Blackberry.

As a result of analysis it was received that iPhone and Samsung are leaders by scores.

One big disadvantage of this feature, however, that Siri should not be able to silent. Even while reducing the voice volume control buttons, the application may not be made silent. Another problem with Siri is obvious that probably come to mind most consumer first: the recognition of voice commands. Such as was the case with many speech recognition programs in the past, virtual assistant can be the problem of recognition of the team, giving a message saying, "unfortunately, I'm having problems connecting to the network."

Siri, although it may do many other things besides just answer the questions. It can set the alarm clock, manage your calendar user, and make phone calls simply by setting it too. A virtual assistant can remind the user about the upcoming meeting, text or call anyone the consumer requests, and to play the role of Navigator on road trips.

Another coincidence with voice recognition that although he understands and speaks three languages, can move only when the consumer commands in English language

The biggest downside to Siri and S Voice is the inability to access voice helper offline.

Both systems had difficulties with the analysis of unique or different spoken words when implementing search in the Internet, distorting enter in the search bar.

The analysis showed that Siri more intelligent than S Voice, and does not require long rehearsals. Besides, the noise reduction system Apple is more perfect.

In the end it will be right to say that the leader is iPhone, but by the cost the leader is Samsung. But anyway both of them, S Voice and Siri are two leaders in this field.

8. Sources

1. BURNETTE, Ed. *Hello, Android: introducing Google's mobile development platform*. 2nd ed. Raleigh, N.C.: Pragmatic Bookshelf, c2009, xvii, 243 p. ISBN 19-343-5649-2.

2. RUSSELL, Stuart J a Peter NORVIG. *Artificial intelligence: a modern approach*. Englewood Cliffs, N.J.: Prentice Hall. ISBN 01-310-3805-2.

3. HARRIS, Michael C. *Artificial intelligence*. New York: Marshall Cavendish Benchmark, 48 p. ISBN 978-160-8700-769.

4. KUMAR. *Artificial Intelligence*. New Delhi: I. K. International Pvt Ltd, 2008. ISBN 9788190656665.

5. PIERACCINI, Roberto. *The voice in the machine: building computers that understand speech*. Cambridge, Mass.: MIT Press. ISBN 02-620-1685-0.

6. ROGOWSKI. Industrially oriented voice control system. 2012, s. 315. Dostupné z: http://dx.doi.org/10.1016/j.rcim.2011.09.010

7. NEUSTEIN, Amy. *Advances in speech recognition: mobile environments, call centers and clinics*. New York: Springer, c2010. ISBN 14-419-5951-3.

8. Using Voice and Speech Recognition with Mobile Devices. [online]. 2013 [cit. 2014-03-17]. Dostupné z: http://www.atarizona.com

9. *Using iPhone's Siri Voice Command*. Indianapolis, Indiana 46240 USA: Pearson Education, Inc, 2011. ISSN 13:978-0-7897-4950-5.

10. BROWN. Top smartphone voice recognition software. [online]. 2013 [cit. 2014-03-17]. Dostupné z: http://www.geeksquad.co.uk/articles/top-smartphone-voice-recognition-software

11. The Pros and Cons of iPhone 4S's Siri. [online]. [cit. 2014-03-17]. Dostupné z: http://www.techcrates.com/iphone-4s-siri-pros-cons/

12. Voice Recognition– Controlling mobile devices using voice commands. [online]. [cit. 2014-03-17]. Dostupné z: http://patseer.com/wp-content/uploads/2013/07/Voice-Recognition-in-Mobile-Devices.pdf