JCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

An Efficient Voice Based Person Identification System for Secured Communication

[1*]Kadher Farook, [2]Manikandan.S, [3]Shakila Basheer, [4]Albert Irudaya Raj

[1]Department of Computer Science, Periyar University, Salem, India.
[2]Department of Computer Science and Engineering, Sriram Engineering College, Chennai, India.
[3]Department of Information System, King Khalid University, Muhayil, Saudi Arabia.
[4] Department of Information Technology and Management, Arul Anandar College, Madurai, India

*Corresponding Author kadherfarook@gmail.com, Tel.: +91-9488226537

Available online at: www.ijcseonline.org

Accepted: 19/Sept/2018, Published: 30/Sept/2018

Abstract — Secured Communication is essential due to scalability due to increase number of devices and drastically growing number of people involved in communication. In this paper a voice comparison based communication authentication mechanism is used for providing secured communication. This voice based authentication is used in two different applications like people communication and data retrieval. Before going to speak with people in online their information and their voice is compared and verified from the database and permission will be granted. Similarly according to the voice they can retrieve the data from the data base, where it provides data integrity. Both applications comprise a number of stages such as: (i) Voice, Voice to Text input, (II). Voice Comparison and Pattern Matching. Finally (III). Permission Granted and Data Retrieval (DR) as the output. In order to improve the accuracy and relevancy the proposed data retrieval system, it uses an indexing method called Bag of Words (BOW). BOW is like an index-table which can be referred to store, compare and retrieve the information speedily and accurately. Index-table utilization in DRS improves the accuracy with minimized computational complexity. The proposed DRS is simulated in DOTNET software and the results are compared with the existing system results in order to evaluate the performance.

***Keywords*:** Information Retrieval System, Data Mining, Bag of Words, Data Base Maintenance.

I. INTRODUCTION

In general IR is an activity is used by a few people for library management, paralegals and the digital library searching system. The world is growing with lots of changes were more than million number of people are using IR in everyday life like email, web searching. After sometime the IR system is used for information access and traditional searching in databases such as, searching an order, searching a product, searching a document from a digital library and so on. It is well known that the IR retrieves data from unstructured databases. The term "unstructured data" means the data is not clear, semantically overt and the format of the data is undefined. Simply can say that it is opposite to structured data (example: DBMS, RDBMS), but in real-time there is no data are not truly unstructured.

Searching information, images, documents and files are created based on the visual appearance and the properties of the data, document and images. Information retrieval is a challenging problem where it has been received a considerable attention from most of the researchers in various fields of image processing, data mining, information retrieval and computer vision and multimedia systems. The growth of web technology brings a drastic increase in data usage published in the recent decades, which has been a great challenge to develop efficient information retrieval systems to help all the users in IR systems. Traditional IR models such as: vector space model [16], classical probabilistic IR models [15] and language modeling approaches [13] are used for query based document retrieval and works independently.

Web search engines are used for entity based retrieval [14, 12] used for commercial purpose. An entity based web document retrieval [9-11] are used in the earlier research works to provide a better semantic based document searching. Searching, information retrieval, content based information retrieval systems are still getting urgent demand in the web applications [17,18] The retrieval system concentrates on features as important for information extraction. Most of the paper follows the feature based IR on content based image retrieval systems [19-21] Some of the IR systems used to transform in order to decompose and represent various resolutions, various sizes and various amounts of information [22-23] Wavelet transform have been successfully applied to image Denoising [24], image compression [25] and texture analysis [26] In [27] the authors propose a new CBIR system using color and texture features. In this paper texture features are extracted Euclidean distance measure to obtain the similarity measurement between a query text and text in the database. In [28] wavelet basis was used to characterize each query image and also to maximize the retrieval performance in a training data set. To make DRS is more efficient, DRS is not constructed based on all the entities. It is query independent. For each voice query the index is selected and then the related data are selected from different location.

One the index is matched, and then DRS decides the location of the data and the entities of the index-data from the database. In this paper the information retrieval system is developed using index searching and pattern matching methodologies. To do index searching BOW is used. The contribution of the proposed DRS work is:

A. Voice (input)
B. Creating BOW
C. Voice Matching and Pattern Matching
D. Communication Permission granted and Data Retrieval (output)

II. RELATED WORK

The proposed model clearly says about the entire functionality of the proposed DRS and it is shown in Figure-1. Any physically challenged people one who are not able to operate the keyboard can use this application. In this paper, it is assumed that the application is developed for online shopping. The user can say about the product in mic then the voice is converted into text. The converted text is taken as a keyword for pattern matching in the product database. During the pattern matching keyword is verified with the BOW in order to check the product availability. If the keyword is available in BOW then the other relevant information about the product is taken from the database, converted into voice, and play back to the user. It is an advanced application can be used in handheld devices also.

In user communication, initially the numbers of users are registered with their voice. The voice is the keyword for comparison, whereas before coming to communicate in online both end user has to be verified by the voice. If the present users' (ready for communication) voice is matched with the stored DB voice then they are permitted for communication and they can proceed. This functionality is depicted in Figure-1.

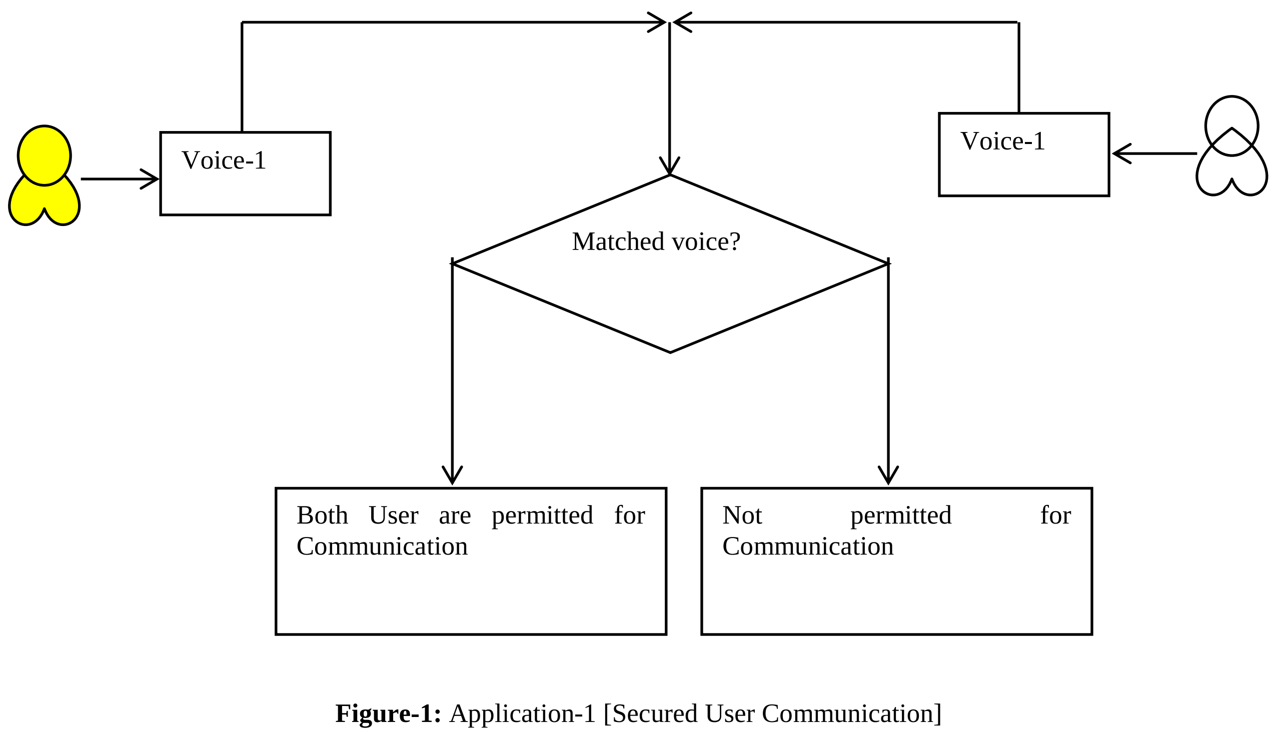

Figure-1: Application-1 [Secured User Communication]

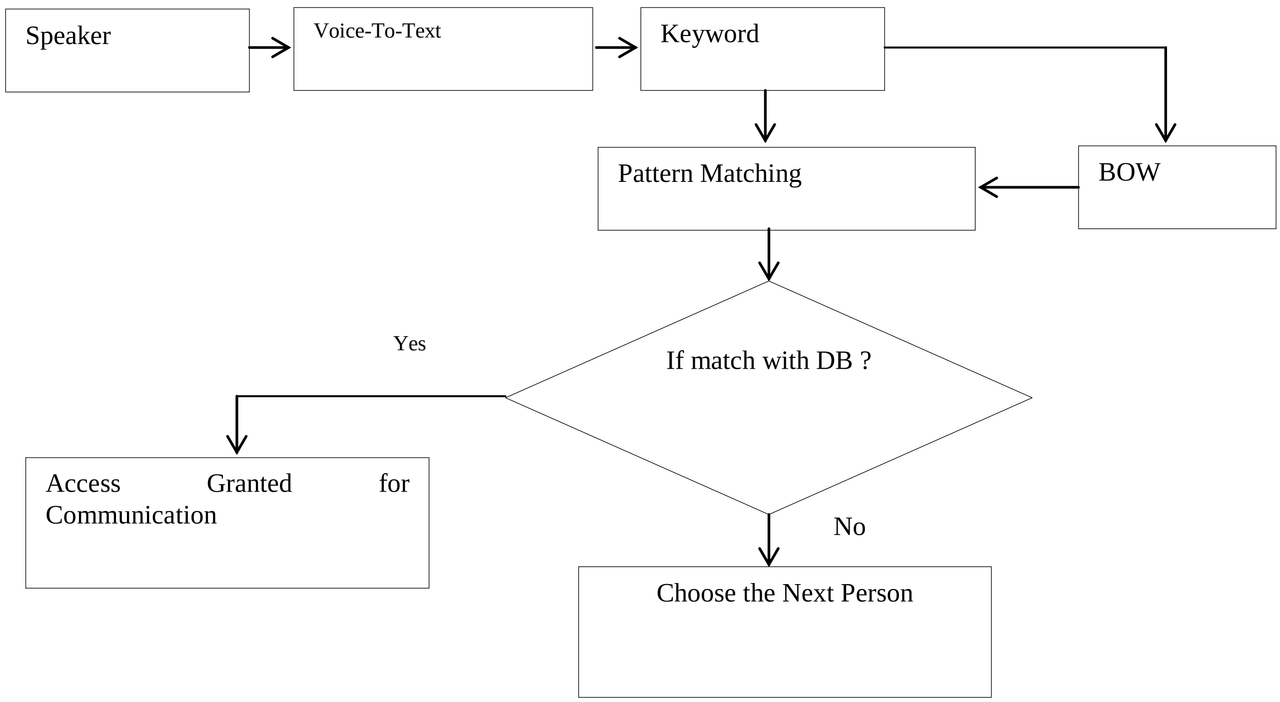

Figure-2: Application-2 [Secured Data Retrieval]

A. Bag-of-Words

One of the most common methodologies to obtain the entire data is by visual words and it can be applied as text indexing and retrieval scheme. The index is created from any one of the feature of the data stored in the DB before persisting newly in the DB. It can be called as Bag-of Words or bag of feature model. Some of the static terms are taken from the data and it is maintained as a catalog (BOW). This catalog is compared with the database data for retrieving the specific data matched with the catalog. The data retrieval using keywords can predict maximum relevant based data and it satisfies the customer.

In this proposed DRS, whenever a new product detail is entered into the database, any one of the data feature is added as index word into the BOW. It needs not be a numeric or character data type and it can arrange the entire BOW automatic while inserting a new index. This automatically arranging of index words helps to compare and retrieve the relevant data speedily and accurately without computational complexity. For example: when a new data d_n is inserted into the database D, one of the feature from the feature set $f_i=\{f_1, f_2, \ldots, f_n\}$ is stored into BOW.

Field-1	Field-2	…….	Field-i	……	Field-n
D_1f-1	D_1f-12	…….	D_1f-i	……	D_1f-n
:	:	:	:	:	:
:	:	:	:	:	:
D_nf-1	D_nf-2	…….	D_nf-i	…….	D_nf-n

BOW
D_1f-i
:
:
D_nf-i

Figure-3: BOW Creation

Each field of the data is considered as separate features and any one of the field is stored into BOW. In an image retrieval system BOW is created automatically using LABELME tool. But in case of alphanumeric data the feature has chosen as keyword manually by the developer according to the convenient. Figure-3 shows the way of BOW creation and it can be used to compare the product availability in the database.

The data classification and retrieval is based on the BOW index, where BOW is the structured features taken from all the trained data inserted in the database. The word stored in the BOW belongs to the same class and it is behaving like a codebook used to cluster and classify the entire dataset. The words of all dictionaries represent frequent structures of all form types. Each word type is represented by a feature vector. The structural features of a form $\mathbf{s_j}$ are calculated and are assigned to the cluster centre $\mathbf{w_i}$ (word) with the smallest (Euclidian) distance mini $\|\mathbf{s_j} - \mathbf{W_i}\|$. This distance is used to fetch the matching BOW for the voice into text (keyword).

B. Voice –To-Text

A portion of the DRS system is programmed to recognize the speech (voice), and convert into text using speech synthesization mechanism available in the system library. The inbuilt speech recognition engine is instantiated initially, and then the defined grammar is loaded in order to recognize the phrases. Adding grammar is used to identify the grammar-name. Each time the grammar is loaded dynamically in order to update the new BOW inserted. This updating can be obtained by the recognizer update method. In this paper the DRS listens to the user whether any speech data is entered into the system. The speech recognition engine is already loaded with the predefined trained text in the background. Each time speech made one line of text is displayed at a time in the system. The main advantage of this system is it will wait for a small interval in order to avoid congestion and proceed with the next BOW. If the speech is understandable by the speech engine then it keeps idle and waits for the next speech and it won't create any software breakup.

The speech to text is an application where it does translate words into text as much as possible due to various countries' accent variation. Other than the DRS, this voice to text conversion is used in healthcare, traffic systems, military, telephony and education systems. It is mainly focused for people with dis-abilities. This paper follows a fuzzy logic based Speech Recognition of Linguistic Content method [1] In this method a word in a language, speaks in different accents, different speeds of pronunciation and with different emphasis. For example, the word "vector" of the English language will be spoken by an American as "vektor", with curtness at the 'c' and at the 't', while a British will speak it as "vectorr", with emphasis on the 'c' and a slight repetition on the 'r'. Similarly, a Russian will speak this word as "vecthor", with softness on the't'. However, the word remains the same, that is, "vector", with slight variations with respect to different accents, speeds of pronunciation and emphasis.

Thus, a single word can be represented as a fuzzy set. However, a word is too specific so as to fit into a generic model of speech recognition. To have a more general model, the fuzzification of phonemes is more appropriate. This model is therefore applied to spoken sentences. One fuzzy set is based on accents, the second on the speeds of pronunciation and the third on emphasis. The use of this method will be especially for speech-to-text conversion, by filtering out the unnecessary paralinguistic information from the spoken sentences.

C. Pattern Matching

In this paper the main idea is to search from right to left in the pattern. With this scheme, searching is faster than average. In order to do this the Boyer-Moore (BM) algorithm positions the pattern over the leftmost characters in the text and attempts to match it from right to left. If no mismatch occurs, then the pattern has been found. Otherwise, the algorithm computes a shift; that is, an amount by which the pattern is moved to the right before a new matching attempt is undertaken. The shift can be computed using two heuristics: the match heuristic and the occurrence heuristic. The *match* heuristic is obtained by noting that when the pattern is moved to the right, it must

1. Match ***all*** the characters previously matched, and

2. To bring a different character to the position in the text that caused the mismatch.

The last condition is mentioned in the Boyer-Moore paper [3], but was introduced into the algorithm by Knuth et al. [2] Following the later reference, we call the original shift table $\boldsymbol{dd}$, and the improved version $\widehat{\boldsymbol{dd}}$. The formal definitions are

$$\widehat{dd}[j] = \min\{s + m - j | s \geq 1 \; and \; ((s \geq i \; or \; pattern[i-s] = pattern[i]) \; for \; j < i \leq m)\}$$

$$\boldsymbol{for \; j = 1, \ldots, m; and}$$

$$\widehat{dd}[j] = \min\{s + m - j | s \geq 1 \; and \; ((s \geq j \; or \; pattern[j-s] \neq pattern[j]) \; and \; ((s \geq i \; or \; pattern[i-s] = pattern[i]) \; for \; j < i \leq m)\}$$

The $\widehat{\boldsymbol{dd}}$ table for the pattern $\boldsymbol{abracadabra}$ is

dd	a	b	r	a	c	a	d	a	b	r	a
$\widehat{\boldsymbol{dd}}$ [j]	17	16	15	14	13	12	11	13	12	4	1

The occurrence heuristic is obtained by noting that we must align the position in the text that caused the mismatch with the first character of the pattern that matches it. Formally calling this table *d*, we have

$$\boldsymbol{d[x] = \min\{s | s = m \text{ or } (0 \leq s < m \text{ and pattern } [m - s] = x)\}}$$

for every symbol $\boldsymbol{x}$ in the alphabet. This methodology is used to compare the voice converted text with BOW and with the database. If the pattern matches the database, then the voice based reply is produced to the physically challenged people. The voice is produced by converting the relevant record information obtained from the database and convert into voice.

D. Text-To-Voice

Text-to-speech synthesis takes place in several steps. The TTS systems get a text as input, which it first must analyze and then transform into a phonetic description. Then in a further step it generates the prosody. From the information now available, it can produce a speech signal. The structure of the text-to-speech synthesizer can be broken down into major modules:

Natural Language Processing (NLP) module: It produces a phonetic transcription of the text read, together with prosody.

• Digital Signal Processing (DSP) module: It transforms the symbolic information it receives from NLP into audible and intelligible speech. The major operations of the NLP module are as follows:

E. Text Analysis:

First the text is segmented into tokens. The token-to-word conversion creates the orthographic form of the token. For the token "Mr" the orthographic form "Mister" is formed by expansion, the token "12" gets the orthographic form "twelve" and "1997" is transformed to "nineteen ninety seven".

F. Application of Pronunciation Rules: After the text analysis has been completed, pronunciation rules can be applied. Letters cannot be transformed 1:1 into phonemes because the correspondence is not always parallel. In certain environments, a single letter can correspond to either no phoneme (for example, "h" in "caught") or several phone me ("m" in "Maximum"). In addition, several letters can correspond to a single phoneme ("ch" in "rich"). There are two strategies to determine pronunciation:

In dictionary-based solution with morphological components, as many morphemes (words) as possible are stored in a dictionary. Full forms are generated by means of inflection, derivation and composition rules. Alternatively, a full form dictionary is used in which all possible word forms are stored. Pronunciation rules determine the pronunciation of words not found in the dictionary.

In a rule based solution, pronunciation rules are generated from the phonological knowledge of dictionaries. Only words whose pronunciation is a complete exception are included in the dictionary. The two applications differ significantly in the size of their dictionaries. The dictionary-based solution is many times larger than the rules-based solution's dictionary of exception. However, dictionary-based solutions can be more exact than rule-based solution if they have a large enough phonetic dictionary available.

Whenever a voice input into DRS it is taken as the query for searching the relevant product from the database. Query enriches expansion is a general strategy used in text retrieval, which is directly adapted to the BOW model in all kinds of data retrieval. In this project the query expansion is simply taken as index searching with BOW and pattern matching with the database. There are various query methods are available like Transitive Closure Expansion (TCE) [4], and Additive Query Expansion (AQE) [5]. In this paper the TCE is used for query processing system. Initially the query word (voice to text) is compared with the index where each visual word has an index indicating that the entire data is available in the database or not. This paper doesn't calculate the score value defining the similarity [6], since the keyword is unique. Using the above text to speech conversion the voice reply is generated and play with the user. The entire functionality of the proposed DRS is given in the form of algorithms, it can be coded in any computer programming language and the efficiency can be evaluated.

```
Algorithm_DRS (string product)
{
Input: voice, product data, initial BOW;
Output: voice
Description:
user speech in mic
Voice is converted into text
Apply a pattern matching algorithm
Search text into BOW
If(text exists in BOW) then search in DB
Voice (" product details"); // all the fields from the matched
field is converted into voice
Else
Voice ("product not available");
End
If any product insertion then
     field-i insert into BOW
}
```

The functionality of the proposed DRS is programmed in DOTNET 2010 software and the results are produced. There are 25 systems are installed in a laboratory in order to evaluate the system performance. In all, the system DOTNET software and the IRS module is installed. The proposed DRSare programmed, experimented in DOTNET software and the results are given below to analyze the performance. One among the systems is assumed as the server and the database is installed. The database is a lexical dictionary which consists of a collection of data in the form of rows. Each row consists of various numbers of columns which is not having appeared like a table. Another system is assumed as a middleware, having BOW table, which consists of a set of all inserted index keywords. Whenever a voice input entry to the system it refers the BOW first and then comes to the database server, which reduces the computational complexity.

In order to experiment the proposed DRS, a product dataset is taken from [8] and experimented. 100 different products are stored in the database. It is assumed that the most of the product names are known by the user and it is online shopping. Some of the product name with some more relevant information about the product is shown in Table-1. Product code, product name are the two main features mostly used for searching the product information speedily in the entire database. Instead of concentrating all modules of online shopping, it is simply coming to know the product availability and product price with other relevant information about the product. The database consists of 15 fields in the table were on our paper only 5 fields are taken as important information to verify the DRS performance. In common product-code is used as searching indexes, but here due to voice mining, product name is used as searching indexes.

Table-1: Product Information

Product Code	Product Name	Product Line	Quantity In Stock	Buy Price	Product Description
S10_1678	1969 Harley Davidson Ultimate Chopper	Motorcycles	7933	48.81	This product is good and u can get world service
S10_1949	Alpine Renault 1300	Classic Cars	7305	98.58	Turnable front wheels; steering function; detailed interior; detailed engine; opening hood; opening trunk; opening doors; and detailed chassis
S10_2016	1996 Moto Guzzi 1100i	Motorcycles	6625	68.99	detailed engine, working steering, working suspension, two leather seats, luggage rack, dual exhaust pipes, small saddle bag located on handle bars, two-tone paint with chrome accents, superior die-cast detail , rotating wheels , working kick stand

There are 100 data is stored in the table where during searching computational time is spent only 100 comparisons and data fetching. For **an N number** of comparisons the computation time taken is **2N+2.** The following figures show that the efficiency of the proposed DRS in terms of accuracy, timeliness and response generations. In order to evaluate the performance, the number of data used in the database table is changed and verified. The number of data is changed from 100 to 1000 and the performance is compared.
In this paper the user provides their input as voice through multimedia input device. The voice is recorded and recognized by the speech engine installed in the system and it is converted into text. The voice recognition is a big process if the Voice-accent is understood by the speech engine then it converts the voice into text. In this process, the number of voices is recognized accurately for the voice input given into the DRS. In order to evaluate the voice recognition accuracy by the DRS, the number of voice input is increased and the recognition rate is calculated. The number of voice input may be changed from 25 to 250. Each round of experiments the number of voice input is increased by 25. Out of the input voice, the number of voices recognized by the DRS system is calculated and shown in Figure-3. Still Google-Voice play is also finding difficulties in terms of voice recognition. In the proposed DRS system the recognition rate is better and it is increased according to the number of voice input increases. The recognition rate is proportionally increased, according to the number of voice inputs getting increased. After successful recognition, the voice is converted into text (it is taken as a keyword) for comparison with the BOW. If the keyword matched with the BOW index, then directly compared with the database in order to process the pattern matching.

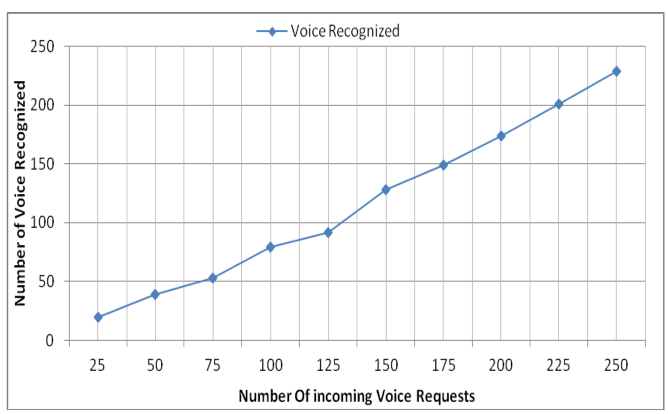

Figure-4: Number of Voice Inputs Recognized Vs. Number of Voice Inputs

If the pattern matched, then the relevant record data are fetched from the data row then converted into voice again. This text-to-voice conversion is played to the user who passed the voice input. According to the number of voice input processed, the number of voice reply is calculated and the quality of the DRS is verified. The number of voice reply against the number of voices is shown in Figure-5. Figure-5 says that the voice reply is increased according to the number of input voice. It is clear that after index matching the reply can be generated according to the pattern availability. The reply may be about the product or it is a message saying that particular product is not available and since there should be a compulsory voice reply for each voice input if it matched with the index. The execution process is preceded when the index is matched, else it is dropped executing the next process. Hence the proposed DRS reduces the computational complexity.

Also Figure-5 shows that the number of voice reply is merely equal to the number voice input given into the proposed DRS. It cannot be concluded that the pattern matching will be performed if the keyword matched with the BOW index due to the product may not be available. The pattern matching algorithms used in this paper find the distance between the possible patterns obtained from the DB with the input pattern. If the distance is merely equal to zero, then the pattern is matched, else it is not matched. According to the pattern matching algorithm, the accuracy is calculated and shown in Figure-6. The percentage of pattern matching is merely equal to the percentage of index matching. From this figure, it is clear that the number of pattern matching is lesser than the number of index matching. After the index matching successful the appropriate pattern may not available in the database and it affects the pattern matching accuracy. It cannot be concluded that the accuracy of the DRS is less. In this paper the accuracy of the entire IR system can be taken as the average of both index matching and pattern matching.

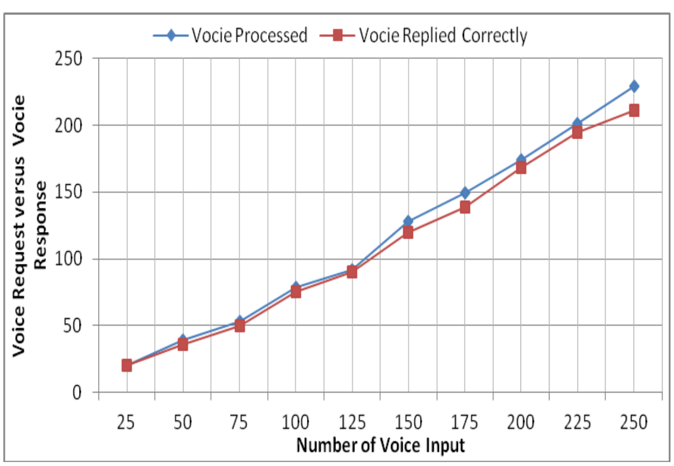

Figure-5: Number of Voice Input vs. Number of Voice Output

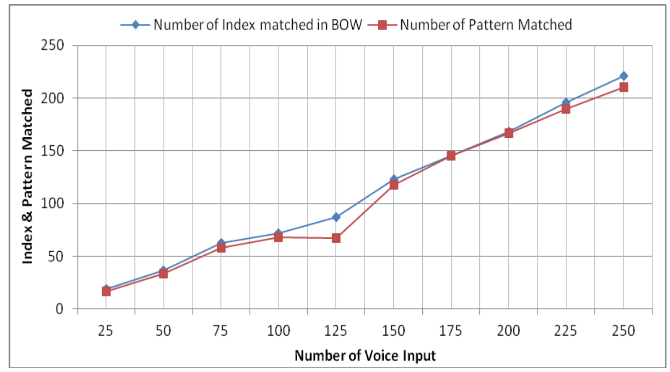

Figure-6: Voice Input Matched with Index and Matched With Pattern Comparison

The computational complexity refers the number of statements in the program to be executed in the compiler and the time taken to compile. The number of statements in the program decides the compilation time and the compilation time taken by the proposed DRS is shown in Figure-7. The figure shows that the computational time is less and it increases, according to the number of inputs increased. It means that for 100 numbers of data it takes only 4 seconds to make the entire process of DRS.

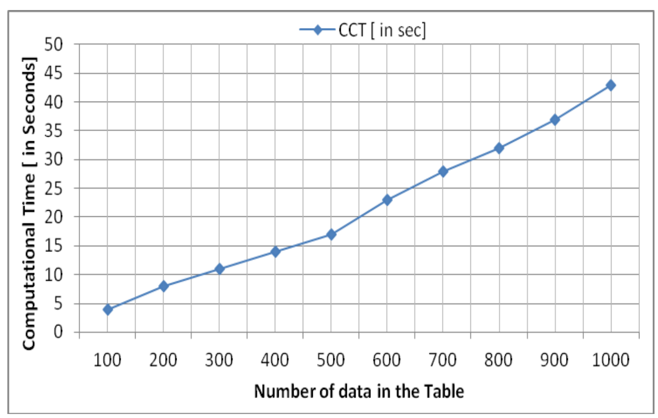

Figure-7: Computational Time In Terms of Data Size

Also the efficiency can be calculated according to the number of response generation against number of input queries. The number of query response against the number of input queries is shown in Figure-9. DRS proved that the number of pattern matching is not depending on the number of index matching completely. It depends on the index matching and the data availability. This figure shows the number of voice reply (response) provided to the user against the query input. The voice reply is gradually increased according to the number voice query applied. The accent and the data availability determine the accuracy of the pattern matching and voice reply accuracy.

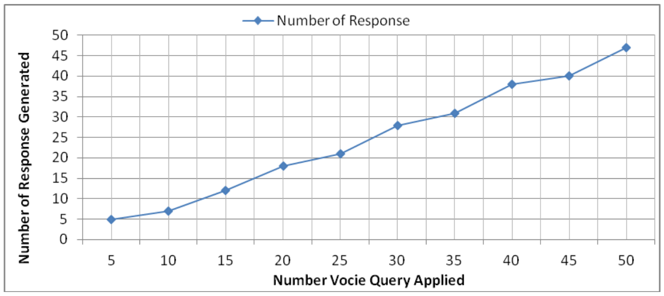

Figure-8: Number Query vs. Number of Response Generated

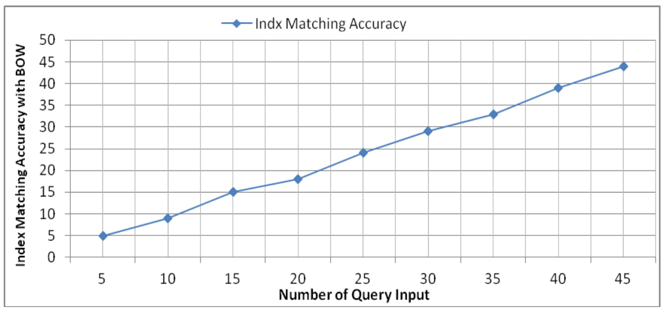

Figure-9: Number of Query vs. Index Matching Accuracy

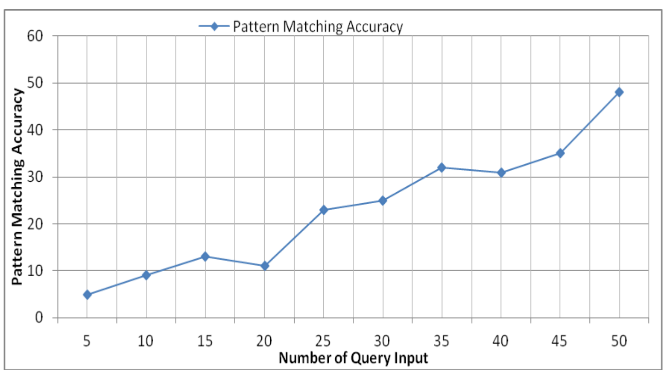

Figure-10: Number of Query vs. Pattern Matching Accuracy

In this paper the number of indices matched and the number of patterns matched is calculated and shown in Figure-9 and Figure-10 respectively. The number of query index matching is proportionally increased, according to the number of query data and accent. The number of pattern matching is up and down in scale due to match pattern and the data available on the DS. In order to evaluate the performance the proposed DRS results are compared with the existing approach.

III. Performance Analysis

The performance of the DRS is evaluated by comparing the mining accuracy and time complexity with the existing approaches [8]. The proposed DRS and the existing IR system are using the data-dictionary at the back end. The data dictionary size is 100, 120 and 140 in terms of number of words. Figure-10 shows the mining accuracy comparison between proposed DRS and the existing IR [8] system. It is clear that the mining accuracy obtained by the proposed DRS is more than the existing IR. To verify the accuracy and comparability the size of the data dictionary is changed gradually and experimented. In each time of the experiment the mining accuracy is also gradually increased in proposed DRS and it is greater than the existing IR accuracy. Time taken to process the query and response generation and for pattern matching is computed for the proposed DRS and compared with the existing IR system. The time taken by the proposed DRS is lesser than the existing approach time. The experiment is repeated for all the dictionary size 100, 120 and 140, and the time calculated. The calculated time includes the voice processing, BOW index matching and pattern matching time. The complete processing time for one job in the proposed DRS is, time from query word is obtained from voice, compared with the BOW, if exists then it compare with the database. Time taken to process the information retrieved by the proposed and existing is shown in Figure-11. From this figure, it is clear that the time taken by the proposed approach is lesser than the existing approach.

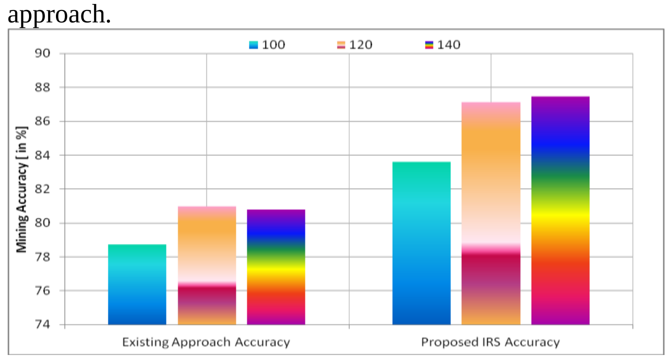

Figure-10: Data Mining Accuracy Comparison between Proposed DRS and Existing Approach

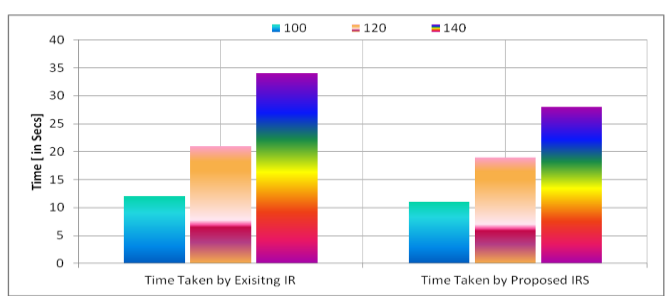

Figure-11: Time Comparison between Proposed DRS and Existing Approach

IV. RUN TIME EFFICIENCY

The efficiency of the proposed DRS is calculated while applying DRS to provide online retrieval and voice reply for large set of database collection. Comparing with the traditional IR approaches, the overhead of proposed voice based DRS comprises four parts: (i). Converting voice to text; (ii). Matching query words in BOW; (iii). Pattern Matching with DB; (iv). Voice based Reply. The previous research off-the-shelf recognition toolkits could already handle the entity annotation on queries well with the high accuracy and low latency. By building BOW using the data features, the overhead of index matching and pattern matching process is reduced to do information retrieval. It reduces the time complexity and computational complexity and since this proposed DRS can be extended to large scale data collection, web applications and in wireless network based applications.

V. CONCLUSION AND FUTURE SCOPE

The main objective of this paper is to develop a data mining model for physically challenged people using voice. The proposed DRS method uses BOW model in order to retrieve the relevant information from the data. Comparing BOW reduces the computational complexity and searching time. In this paper the proposed DRS handle a smart way of information retrieval approach, which estimate the data availability by comparing the index in order to reduce the time and computational complexity. It can be applied for high – dimensional data entity space. This proposed DRS provides voice to text, text to voice and visual word comparison for improving the efficiency of the information retrieval system. From the results it is clear that this approach is efficient in term of reduced computation complexity, reduced time and it is a special kind of information retrieval system helps to social for physically challenge people like blind and no able to operate keyboard. This voice comparison based authentication can be utilized in various kinds of applications and it is proved.

REFERENCES

[1] Lakra, Sachin, et al. "Application of fuzzy mathematics to speech-to-text conversion by elimination of paralinguistic content." arXiv preprint arXiv: 1209.4535 (2012).
[2] Kunth, D. J. Morris and V.Partt.1977. "Fast Pattern Matching in Strings" SIAM J on Computing, 6, 323-50.
[3] Boyer R., and S Moore. 1977. "A Fast String Searching Algorithm." CACM, 20, 762-72.
[4] Ondrej Chum, James Philbin, Josef Sivic, Michael Isard, and Andrew Zisserman. Total recall: Automatic query expansion with a generative feature model for object retrieval. In ICCV, pages 1–8, 2007.
[5] HHerv´eJ´egou, Matthijs Douze, and Cordelia Schmid. Improving bag-of-features for large scale image search. International Journal of Computer Vision, 87(3):316–336, 2010.
[6] James Philbin, Ondrej Chum, Michael Isard, Josef Sivic, and Andrew Zisserman. Object retrieval with large vocabularies and fast spatial matching in CVPR, 2007.
[7] http://www.mysqltutorial.org/mysql-sample-database.aspx.
[8] Kleber, Florian, Markus Diem, and Robert Sablatnig, "Form classification and retrieval using bag of words with shape features of line structures"-IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics, 2013.
[9] M. Banko, M. J. Cafarella, S. Soderland, M. Broadhead, and O. Etzioni. Open Information Extraction from the Web. In IJCAI, volume 7, pages 2670–2676, 2007.
[10] M. J. Cafarella, J. Madhavan, and A. Halevy. Web-Scale Extraction of Structured Data. ACM SIGMOD Record, 37(4):55–61, 2009.
[11] S. Cucerzan. Large-Scale Named Entity Disambiguation Based on Wikipedia Data. In EMNLP-CoNLL, volume 7, pages 708–716, 2007.
[12] T. Lin, P. Pantel, M. Gamon, A. Kannan, and A. Fuxman. Active Objects: Actions for Entity-Centric Search. In WWW, pages 589–598, 2012.
[13] J. M. Ponte and W. B. Croft. A Language Modelling Approach to Information Retrieval. In SIGIR, pages 275–281, 1998.
[14] J. Pound, P. Mika, and H. Zaragoza. Ad-hoc Object Retrieval in the Web of Data. In WWW, pages 771–780, 2010.
[15] S. E. Robertson and S. Walker. Some Simple Effective Approximations to the 2-Poisson Model for Probabilistic Weighted Retrieval. In SIGIR, pages 232–241, 1994.
[16] G. Salton, A. Wong, and C.-S. Yang. A Vector Space Model for Automatic Indexing Communications of the ACM, 18(11):613–620, 1975.
[17] Kherfi, M.L., Ziou, D. and Bernardi, A. (2004) Image Retrieval from the World Wide Web: Issues, Techniques, and Systems. ACM Computing Surveys, 36, 35-67. http://dx.doi.org/10.1145/1013208.1013210
[18] Datta, R., Joshi, D., Li, J. and Wang, J.Z. (2008) Image Retrieval: Ideas, Influences, and Trends of the NEW Age.ACM Computing Surveys, 40, 1-60.
[19] Yang, M., Kpalma, K. and Ronsin, J. (2010) A Survey of Shape Feature Extraction Techniques. Pattern Recognition, 1-38.
[20] Penatti Otavio, A.B., Valle, E. and Torres, R.da.S. (2012) Comparative Study of Global Color and Texture Descriptors for Web Image Retrieval. Int. J. Via. Commun. Image R, 359-380.
[21] Deselaers, T., Keysers, D. and Ney, H. (2008) Features for Image Retrieval: An Experimental Comparison. Information Retrieval, 11, 77-107.
[22] Mallat, S.G. (1989) A Theory for Multi resolution Signal Decomposition: The Wavelet Representation. IEEE Transaction Pattern Analysis and Machine Intelligence, 11, 674-693.
[23] Sarck, J.L., Murtagh, F.D. and Bijaoui, A. (1998) Image Processing and Data Analysis: The Multi scale Approach.
[24] Hill, P., Achim, A. and Bull, D. (2012) The Undecimated Dual Tree Complex Wavelet Transform and Its Application to Bivariate Image Denoising Using a Cauchy Model. 19th IEEE International Conference on Image Processing (ICIP), 1205-1208. http://dx.doi.org/10.1109/icip.2012.6467082.
[25] Kalra, M. and Ghosh, D. (2012) Image Compression Using Wavelet Based Compressed Sensing and Vector Quantization. IEEE 11th International Conference on Signal Processing (ICSP), 1, 640-645.
[26] Kokareh, M., Biswas, P.K. and Chatterji, B.N. (2005) Texture Image Retrieval Using New Rotated Complex Wavelet Filters. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 35, 1168-1178.
[27] Balamurugan, V. and Anandha Kumar, P. (2008) An Integrated Color and Texture Feature Based Framework for Content Based Image Retrieval Using 2D Wavelet Transform. IEEE International Conference on Computing, Communication and Networking, 1-16. http://dx.doi.org/10.1109/icccnet.2008.4787734
[28] Quellec, G., Lamard, M., Cazuguel, G., Cochener, B. and Roux, C. (2012) Fast Wavelet-Based Image Characterization for Highly Adaptive Image Retrieval. IEEE Transactions on Image Processing, 21, 1613-1623.

