Journal of Network and Innovative Computing
ISSN 2160-2174 Volume 4 (2016) pp. 264-271

Semantic Similarities in Voice Information Retrieval System for Documents

Olufade F.W Onifade[1], Ayodeji O.J. Ibitoye[2]

[1]Department of Computer Science, University of Ibadan, Nigeria.
fadowilly@yahoo.com

[2]Department of Computer Science and Information Technology, Bowen University, Nigeria.
ibitoye_ayodeji@yahoo.com

Abstract: **Recent advances in text to speech and vice versa have opened a new dimension to the manner via which information is sought on the go. It is becoming a common place for smart device users to search by voice. This development is driving a new order in Information Retrieval circle with researcher trying to improve on the design of information retrieval systems operated by voice. Techniques like the spoken text detected, word transcription, phonetic query expansion in voice information retrieval system for document tend not to retrieve relevant documents to users speech data due to the problems of polysemy, query drift, low precision and low recall values engendered from clustered document. Therefore, this research extends the conceptual query model by using the new concept (term) generated by the Universal Fuzzy Concept Network Language with Latent Semantic Analysis technique to build semantic similarity between concepts in a document before retrieval. Also, the research find the degree of relational relevance concepts in potential retrievable document. Here, the research achieved a higher degree of cohesion, lower degree of entropy between concepts and better semantic similarity of concepts in potential retrievable documents to the user query. When compared to keyword spotting and query expansion technique, a better precision and recall values was also achieved by using the proposed method**.

Keywords: Latent Sematic Analysis, Universal Network Language, Fuzzy Concept Network, Conceptual Structure, Voice Information Retrieval System for Documents, Information Retrieval, Concept.

1. Introduction

Information retrieval (IR) is a process by which a user finds and retrieves information, relevant to the user, from a large store of information. The goal of IR is to retrieve all of the information a user needs and at the same time limit the irrelevant information that is retrieved for the user [1]. Since the informational seekers are not aware of the relevant information that is available to their query, a just-in time information retrieval agent [2] has been used to retrieve and present in sequence relevant information that is based on users local situation in an easily accessible but yet nonintrusive manner. However, several problems associated with information retrieval has perpetually become larger since the web databases which are accessible to users for information seeking activities have become larger in volume, faster in growth and enhanced with more effective technologies to upload variety of web content from different sources [3]. In obtaining a relevant or near relevant information need, users are required to communicate their thought pattern to the information retrieval system (IRS) effectively. Communication as a process of sending and exchanging information is a core requirement whose output either in speaking or writing determines the input either through listening or reading. For communication to take place effectively especially in information seeking process, at least two systems must be put in place. One of such system must put something "out" while the other system must take something "in". This is called “output” and "input" processes [4]. The “out”, “in” scenario of communication best describe what transpired between an information seeker who outputs his/her intention and an information retrieval system which take in the query for processing. One significant problem a user can encounter in presenting a need to an IRS is the problem of Translating needs to request and request to query. Translation here means representing or presenting a thought (need) in the most precise and accurate way as a request for an IRS to understand [5]. Figure 1 is used to further illustrate the problems encountered by users when presenting a need to an IRS.

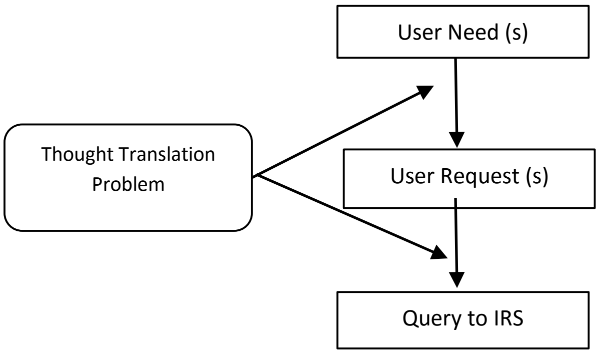

Figure 1: Query Definition and the IRS

Majority of the queries; which are formal expressions of information needs, for example search strings in web search engines have gone beyond typing constructive request in a text field to communicating with an information retrieval system through voice. Hence, for an effective result to be obtained by

IR systems, effective communication must take place between the searcher and the information retrieval system from the start of problem definition to the end of interpreting the retrieved document [6]. It is import to note that in information retrieval a query does not uniquely identify a single object in the collection. Instead, several objects may match the query, perhaps with different degrees of relevancy. Over the years, information retrieval systems via text have played a long vital role in the structuring, analysis, organization, storage, searching and retrieval of information [7]. However, her values have not been fully optimised since web retrieval is made increasingly difficult when adding in factors such as word ambiguity (where a single word can take on multiple meanings), and the large amount of typographical errors contained within web information. Currently, voice is becoming more and more a preferred medium of interaction between people and the World Wide Web [8]. However, due to existing technical difficulties and limitations with natural language processing, grammar generation, voice recognition, result representation, amongst others, a very little attention has been given to document information retrieval using voice. Voice search is the technology intended to provide users with the information they request with a spoken query [9]. In this research, we used the new queries generated by establishing conceptual relationship in spoken text detected from user's speech with the Latent Sematic Analysis (LSA) to build semantic similarities with the view to increase the percentage of retrieving more relevant documents in respect to user's queries.

In the remaining sections of this work, we examined related works in the realm of VIRD and their qualities in section 2. Section 3 gave an overview of latent semantic analysis. In section 4, we discussed the Universal Fuzzy Concept Network Language (UFCNL) by detailing the processes of the fuzzy concept network language and the roles of universal network language in generating new concept. Also, the effect of LSA in spanning relevant document to users was also discussed in section 5. In section 6, we presented the experiments by using the model in 4 to evaluate 20 distinct user queries on a clustered data set of 3200 to test for the precision rate, recall rate and degree of cohesion to compare existing methods like keyword spotting, query expansion and UFCNL before concluding with the research documentation in section 7.

2. Voice Information Retrieval for Document

Since, human beings finds it easier to communicate and express their ideas via speech [10], the introduction of speech language technologies into Information Retrieval systems has enhanced a more natural communication between IR systems and users in need of information. The voice search system which now presents a graphic user interface for IR has the capacity to transcribe spoken queries and use the text output for information search. In voice information retrieval for document, query is given as voice command to the system; the system has to generate the information based on the spoken word to retrieve the needed documents [11]. Several methods have been used to implement a voice information retrieval system for documents (VIRD). From the telephony system of interaction with the indexed database to VoiceXML and to using a large vocabulary continuous speech recognition (LVCSR) [12] are approaches that help by using its language information to search for contents that include speech in the database. However, the lack of sufficient training data to produce high accuracy transcript, identification and retrieval of queries in speech data from low-resources languages remains a major challenge [13]. A more recent approach for VIRD is the Automatic Speech Recognition (ASR) mechanism which is used together with language processing for the purpose of producing transcripts of spoken words which are useful for document retrieval [14]. The ASR system typically employs a closed-vocabulary approach which calls for a predetermination of speech. Therefore, in building an ASR system for an information retrieval application, the choice of words in the system's vocabulary system is vital. Although it has been stated [15] that transcription errors greatly influence the performance of voice search systems. The goal of using a probabilistic model like the hidden Markov model (HMM) is to predict the most likely string of words given the observed information. Thus the HMM among other, statistically represent speech event like a word, wherein, model parameters are typically trained on large corpus of speech data. The ASR begins it task by firstly acquiring the human voice, converting to proper format and storing the data. Thereafter, a number of predefined features such as noise will be extracted from the process speech signal [16] before a pluggable module which contains the Language Model, the Dictionary, and the Acoustic Model, and also allows users to dynamically configure the system with different Linguist implementations will be used to generate Search Graph that is used by the decoder during the search [17]. Here, the Search Graph is the primary data structure used during the decoding process. Finally, the Decoder, which use Features from the Front End in conjunction with the Search Graph from the Linguist to generate Result hypotheses [18]. Even though transcription errors greatly influence the performance of voice search systems. The Recent research has vastly improved the transcription quality of ASR frontend of search engines, as well as their handling of verbosity in ranking [19]. Hence, ASR approach seems good because speech recognition allows the user to query the system by voice, making the interaction direct, easy to use and more immediate. Also, [20] introduced an approach to detecting speaker's state based on acoustic, prosodic and phonotactic features. These features are independent from the ASR outputs and can be extracted from the audio itself. The acoustic features extract the pitch contour of the of the posting list entry and compute its median, as well as the pulses, duration, jitter and shimmer while the prosodic features automatically detect and classify prosodic events like pitch accents and prosodic phrase boundaries. This plays an important role in capturing many aspects of the manner in which words are spoken. However, unlike text retrieval, the basic challenge of VIRD is dealing with speech recognition errors and out-of-vocabulary (OOV) words document processing [21]. OOV terms are words missing in the automatic speech recognition (ASR) system's vocabulary and are replaced in the output transcript by alternative words which are probable, given the language models of the ASR. With the OOV present in query search, no retrieved document will be relevant to users query. In resolving the OOV problem, several IR technique of query expansion have also been introduced to tackle OOV. Some of this technique involves finding synonyms of query terms from the speech to finding all the various morphological and adding additional words by using techniques such as relevance feedback [22]. Many other studies have also focused on dealing with OOV and misrecognized words. Such of this methods include the example of using sub word units as the recognition result or search unit by using the multiple recognition candidates [23]. One other approach for solving the OOV issue consists of converting the speech to phonetic transcripts and representing the query as a sequence of phones. Tomoko [24] also proposed a method that incorporates Spoken Text Document (STD) into Spoken Content Retrieval (SCR) process to deal with OOV and miss-recognized words. However, reviewed VIRD approach mostly use keyword search from speech recognition, the option give existing VIRD systems a low precision and recall value [25]. Also, the level of accuracy in speech transcription, particularly on terms of interest such as named entities and content words are issues with retrieving information from spoken data [26]. Here, the research proposed a formal language approach to be used alongside the fuzzy concept network technique for generating and establishing a User-IR model that engenders conceptual relationship between the user queries before search is conducted. Thereafter, Latent Semantic Analysis technique is used to build semantic similarity structures in potential retrievable document where the output will be documents with high cohesion rate, and lower entropy in their respective ranking order. In section 3, we presented an overview of Latent Semantic Analysis and its role in establishing semantic similarity structure in documents

3. An Overview of Latent Semantic Analysis

Latent semantic indexing is a retrieval technique that indexes and uses a mathematical technique called Singular Value Decomposition (SVD). It does not only identifies pattern in an unstructured collection of text, it also find relationship between patterns [27]. Since, most approaches to retrieving textual materials depend on a lexical match between words in users' requests and those in or assigned to database objects. LSA allows one to find more relationships between terms and to get a more differentiated view on the data set and its underlying relations. There have been proposed number of approaches for automatic discovering of semantic relations between words: with help of lexical and dependency patterns [28], based on Latent Semantic Analysis from evidence contained in electronic dictionaries or encyclopedias, and even from the Web link structure [29]. The tremendous diversity in the words information seekers use to describe the same object, lexical matching methods are necessarily incomplete and imprecise [30]. Hence, the latent semantic indexing approach tries to overcome these problems by automatically organizing text objects into a semantic structure more appropriate for matching user requests. Thus, a well-known method for improving the quality of similarity search in text is called LSI in which the data is transformed into a new concept space.

4. Conceptual Querying Support Model

Concept-based IR represents both documents and queries using semantic concepts in place of keywords extracted from user query in a concept space. The essence is to establish a conceptual knowledge of the user's query. Hence, the conceptual query support model is designed as an information retrieval system that allows users to communicate their intentions through voice command. The model is designed using the UNL and FCN approaches. These methods act dynamic roles for the purpose of minimizing noise, and increasing the level of cohesion by finding and establishing the conceptual structures for a speech based user query before performing a search. This system is trained to listen through voice communication on a speech recognition device that can identify words and convert the user's speech to text format that is needed to build a conceptual relationship for the captured query. Here, the Universal Fuzzy Concept Network Language is used for the generation of conceptual query vectors by utilizing the conceptual relations and Universal Words that are extracted from the user speech to describe the objectivity of the captured sentences [31]. Figure 2 is a detailed diagram that illustrates the flow of the proposed UFCNL.

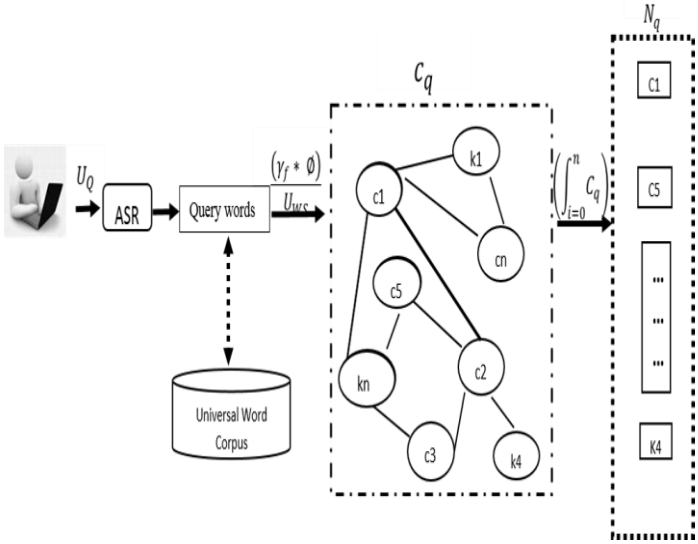

Figure 2: Diagram illustrating the conceptual structure of the user query

From figure 2, U_Q is the user query, U_{Ws} is the Universal word corpus, γ_f is the universal network language, $\emptyset$ is the fuzzy concept network, C_q is the conceptual query and N_q is the new set of ranked concepts from the conceptual queries.

The Universal Network Language γ_f in figure 2 is used on the captured text in the user speech U_Q in order to establish a conceptual relations between function words and concepts. γ_f is a digital meta-language for describing, summarizing, refining, storing and disseminating information in a machine independent and human language neutral form. Since UNL is used as a language for knowledge representation in Information Retrieval, here, the meta-language focuses to express meanings in a group of word retrieved by the ASR. In UFCNL, the Universal Words (UWs) are the main concepts or words that constitute the vocabulary which are always represented by a node in a hypergraph while function words such as auxiliaries and determiners are attributes to UWs that provides additional information to the sentence These words are loaned from English and disambiguated by their positioning in a knowledge base of conceptual hierarchies. Hence, UFCNL uses the UNL to establish relationship between the auxiliaries or/and determiners and the available concepts in the user query in order to define the semantic links wherein high-level concepts can be related to lower-level ones through the ontological relations "equ" (= is equal to), "icl" (= is a kind of), and "iof" (= is an instance of), logical relations ("and", "or") and thermal relations; such as "tim" = time, "plc" = place, etc.

In line with the objective to increase the conceptual cohesion between concepts, the Fuzzy concept network $\emptyset$ is used in UFCNL to build relationships and increase the degree of cohesion between concepts being multiple defined in order to extend the generality of the knowledge base architecture [32]. It utilizes the universal word corpus to generate and build a wide range of concept with fuzzy relative degree of association. In the UFCNL, FCN is modeled by a relation matrix and a relevance matrix, where the elements in a relation matrix represent the fuzzy relationships between concepts, and the elements in a relevance matrix indicate the degrees of relevance between concepts. Hence, we present a relational relevance matrix in equation 1

$$R_m = \begin{array}{c} \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{array} \left\{ \begin{array}{cccc} c_1 & c_2 \ldots\ldots & c_n \\ u_{11} & u_{12} \ldots\ldots & u_{1n} \\ u_{21} & u_{22} \ldots\ldots & u_{2n} \\ \vdots & \vdots & \vdots \\ u_{n1} & u_{n2} \ldots\ldots & u_{nn} \end{array} \right\} \quad (1)$$

Where the element u_{ij} represents the relevance degree between concepts c_i and concepts c_j when they are connected by fuzzy relationship r in some contexts, where $r \in \{P, N, G, S\}$, $u_{ij} \in [0,1]$ and $0 \leq i, j \leq 1$. If the fuzzy relationship r is symmetric, then $u\left(c_i, c_j\right) = u\left(c_j, c_i\right)$. If the fuzzy relationship r is reflective, the $u\left(c_i, c_j\right) = 1$. Otherwise $u\left(c_i, c_j\right) = 0$ when this happens, concepts c_i and concepts c_j are not relevant by any fuzzy relationship.

The concepts has presented in the matrix are explicitly linked by one of the four fuzzy relationships at a time, i.e., fuzzy positive (P), negative (N), generalization (G) and fuzzy specialization (S) association relationship. The essence is to provide a more powerful knowledge representation method which is more appropriate than the Universal Network language in the information retrieval environment. Instead of establishing binary relation between concepts, based on available concepts in a sentence (user query), FCN is used to broaden the scope of the universal words by utilizing the available words present in the Universal words corpus as illustrated in figure 1. Then fuzzy relations with degree of relationship within the range of 0 and 1 are established for each pair of concept. Hence the FCN in UFCNL allows the users to perform positive queries, negative queries, generalization queries, and specialization queries in vector space. From this, the new conceptual query C_q are rank based on degree value of association to form the new set of queries required to perform a search using the function$\left(\int_{i=0}^{n} D_r\right)$.

5. Building Semantic Similarities in Document

Latent semantic analysis is applied on the new generated conceptual query and the clustered document representation to perform a search. At first, each indexed document is converted into a vector of word occurrences. Thereafter, each vector is scaled so that every concept reflects the frequency of its occurrence in context. The column vectors in a large concept-document matrix is then combined with the rows. Here the rows represent concept, while, columns represent documents. Then SVD is performed on the concept-document matrix. The essence is to output documents with the most related conceptual relationship. Figure 3 below further illustrate this.

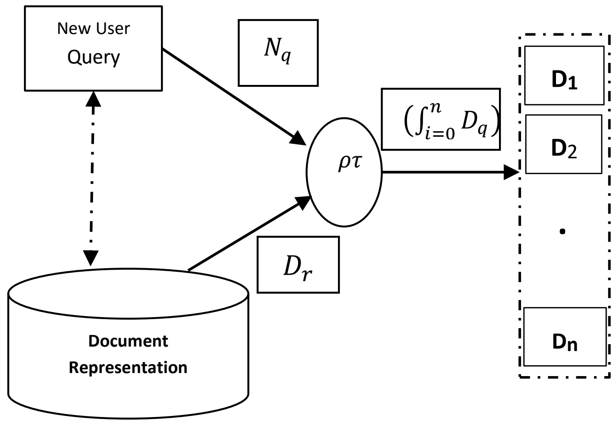

Figure 3: Document Ranking through Semantic Similarities

From figure 3, N_q are new set of concepts generated form the user query through the UFCNL, $\rho\tau$ is the Latent Sematic Analysis while D_r is the represented document in correspondence to the users query.

We used three sub-matrices T_k, S_k and D_k to describe the coordinate system, the semantic space for a document collection. The derived concepts or topics of the document collection are depicted in D_k, and the word distribution patterns in T_k. In spatial terms, the rows of the matrices T_k and D_k are the coordinates of points representing the concepts and documents in reduced k dimensional space. The matrix S_k is used to rescale the axes in order to be able to compare different objects to each other. Here we first of all calculate the concept by document relationship using the vector of the matrices

$$CD_K = D_K S_K^{\frac{1}{2}} \qquad (2)$$

before we find the relationship between the concepts in the documents using the distance in their vector matrices

$$CD_K = D_K S_K \qquad (3)$$

to calculate their cosine similarities. After such decomposition, we when straightforward to recompose and combine Concept Document DC matrix in other to compute a particular matrix M, that can identify the most related concept document with relevance to the user query. Here we have

$$M = DC^t * T_k * S_k^+ * D_k \quad (4)$$

Where S_k^+ is the inverse of S_k.

In the process of using LSA with the new conceptual query and document representation, the model pay much attention to representing conceptual information in clustered document.

Queries	Precision Value			Recall Value		
	Keyword Spotting	Query Expansion	UFCNL	Keyword Spotting	Query Expansion	UFCNL
Q_1	45.53	53.54	59.54	62.44	69.44	74.17
Q_2	43.88	54.28	60.46	60.68	68.51	72.45
Q_3	43.65	53.17	59.68	60.78	68.12	74.76
Q_4	45.83	53.34	60.74	60.74	69.23	75.62
Q_5	46.68	50.45	58.72	63.98	69.45	75.67
Q_6	46.81	51.26	58.16	61.72	70.52	76.75
Q_7	44.41	52.33	60.14	60.49	69.1	75.21
Q_8	46.55	52.41	59.24	59.14	70.34	76.43
Q_9	45.42	51.34	59.86	61.62	69.35	75.66
Q_{10}	45.46	50.88	61.88	62.85	69.62	75.39
Q_{11}	42.83	52.45	58.46	60.24	68.42	73.67
Q_{12}	44.42	53.34	58.92	61.66	68.41	74.77
Q_{13}	44.29	50.81	59.21	62.47	69.43	74.82
Q_{14}	43.67	50.79	58.44	60.25	69.72	73.65
Q_{15}	44.01	52.45	59.23	61.44	68.25	73.82
Q_{16}	43.84	53.14	58.76	61.37	68.49	74.25
Q_{17}	45.56	52.67	58.46	62.74	69.33	75.33
Q_{18}	46.02	53.37	59.23	60.01	68.36	74.91
Q_{19}	46.44	52.6	59.76	60.23	68.78	74.83
Q_{20}	42.78	51.46	58.48	61.97	69.24	74.87

Users would prefer to retrieve documents of interest without having to define the rules for their spoken queries. From the successful construction of the concept by document matrixes, the next approach was to decompose these matrixes into subs using the standard value decomposition. During this process the concept by document level of association that existed therein were established by applying the cosine functions. The cosine function for two vectors of a frequency matrix in standard vector space can only return values in the range [0; 1], because in the traditional vector space model (VSM) all vectors lie within the positive quadrant. The values in any frequency matrix are always positive and a negative frequency count is not possible. A cosine of 0 for two vectors in standard vector space means that their distance is maximal which implies that they are maximally dissimilar. In contrast to VSM, values in a latent semantic analysis (LSA) matrix can be negative so the cosine of two vectors in a LSA space can have a value anywhere in the range of [-1; 1]. Thus we find more relationships between concepts in the document while LSA on concept based queries was able to retrieve documents that did not even share a single word with the query but were rather semantically related. In the next section, we documented the experiment conducted and the result obtained using the proposed methodology on sampled queries over a set of clustered documents.

6. Experiments and Evaluation

The experiment was conducted to reveal the cohesion rate, entropy rate, precision and recall rate in percentage increase by which the proposed technique outperforms existing approaches like keyword spotting and query expansion. The work examined 20 distinct spoken queries on a dataset of 3200 clustered documents to also find the mean average precision, mean average recall, and mean average cohesion and entropy values for user speech data in correspondence to the represented document. In calculating the precision and recall rate for the user queries, we use the equations 5 and 6 as illustrated below:

$$Precision\ rate = \frac{R_{rd}}{R_{dr}} \quad (5)$$

$$Recall\ rate = \frac{R_{rd}}{R_e} \quad (6)$$

While R_{rd} denotes the number of relevant retrieved document, R_{dr} denotes the number of retrieved document and connotes R_e the number of relevant documents in the collection. The figures below reflect the precision and recall rate of 20 different queries as presented in table 1

Table 1: Precision and Recall Values for twenty user queries

In calculating the cohesion and entropy rate using document clustered from table 1 queries, we applied the formula

$$E_J = -\sum P_{ij}\log(P_{ij}) \quad (7)$$

to perform the entropy test. The essence is to measure the quality of document with the caveat that the best entropy is obtained when each document contains exact relevant concept without or with minimal inner noise (irrelevant terms). For each document, the concept distribution of the query is calculated first, i.e., for document *j* we compute *pij*, the "probability" that a member of concept *i* belongs to document *j*. In computing the similarity between concepts, we define

$$c = \frac{1}{S}\sum_{d\in S} d \quad (8)$$

i.e the vector obtained by averaging the weights of the various terms that are present in a document d from the set of documents *S*. Figure 4 and 5 is used to further illustrate the obtained result.

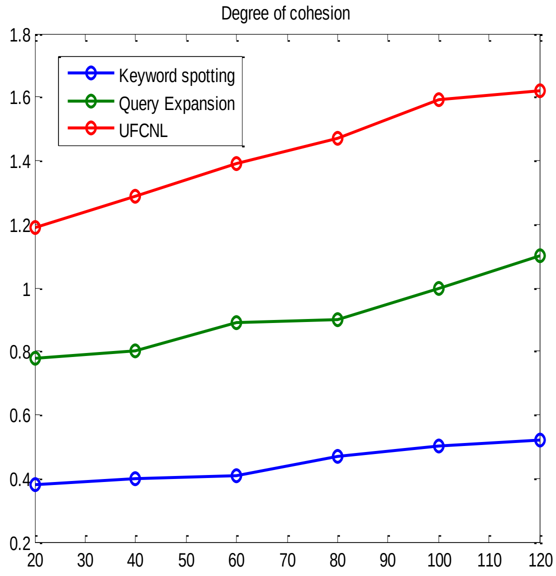

Figure 4: Degree of Cohesion

From figure 5, the degree of cohesion between concepts in potential retrievable documents is a similarity test to measures the level of meaningful association between concepts in documents. Based on information retrieval system condition, the minimal the number of terms in a document, the lower the tendency for query drift and the higher the cohesive force between terms in such documents. While considering keyword spotting, query expansion and the proposed UFCNL, the degree of cohesion in the three technique trend is proportional to the size of document, however, UFCNL achieved a higher degree of cohesion between concepts in documents when compared to the other two approaches at all levels.

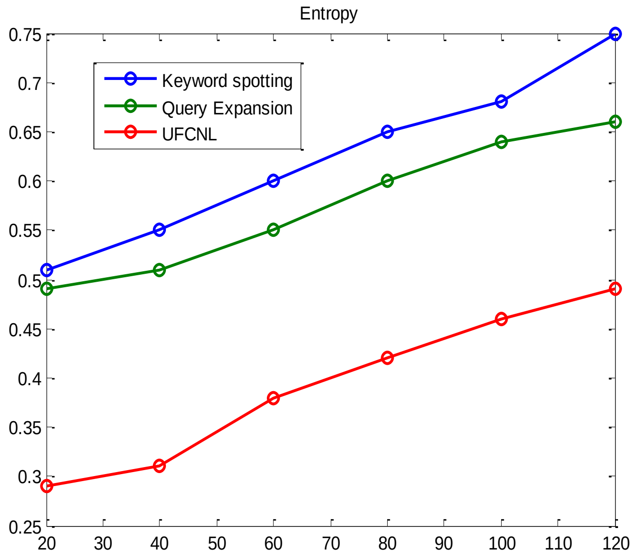

Figure 5: Entropy value

In figure 6, we presented the entropy result from user query using the clustered dataset. Since, a document with a higher number of concept tends to also have a higher level of inner noise (irrelevant terms) when compared to documents will fewer number of concepts. Over the range of document sizes clustered from user's queries, minimal entropy value was obtained for the proposed UFCNL when compared to query expansion alongside keyword spotting.

In addition, from table 1, we evaluated the mean average recall (MAR) value and mean average precision (MAP) value. The MAR tested for the average fraction of relevant documents that are retrieved by the system out of the entire document collection, while the MAP tested for average fraction of documents retrieved by the system that are relevant. In the precision test, the proposed technique outperformed query expansion and keyword spotting algorithms by 7% and 14.5% respectively. While evaluating the recall values, UFCNL engendered 5.8% relevant document when compared to query expansion and 13.5% when compared to keyword spotting algorithm. In like manner, we also tested for the mean average degree of cohesion that existed in potential retrievable document clustered by the user query alongside the mean average entropy ratio. The cohesive test is necessary to evaluate the degree to which concepts are related in a set of clustered documents as compared to the traditional keyword spotting approach, the query expansion technique. By using a threshold value of 0.5, the degree of relationship between clustered words in UFCL approach is 23.5% and 30.2% better to query expansion and keyword spotting approach respective. The inner noise value is also minimal with the new technique as compared with the existing solutions. Here, keyword spotting and query expansion are 24.8% and 8.2% higher than the new technique. From all findings in this experiment, the obtained results indicates that the proposed technique for voice information retrieval system for document outperforms the existing keyword spotting and Query expansion technique in all experiment conducted. This approach in no doubt has helped to enhance the retrieval of relevant document from the data corpus based on the users query in voice information retrieval system for documents. The above data are present in table 2 below.

Table 2: Summarized Experiment and Result

Approach	MAP (%)	MAR (%)	MAC (%)	MAE (%)
Keywords spotting	44.91	61.34	54.42	75.65
Query Expansion	52.30	69. 10	61.22	67.01
UFCNL	59.37	74.85	84.6	50.823

7. Conclusion

The need to retrieving relevant document based on users' speech queries was the main given contribution of this research. At first, establishing conceptual knowledge of the user's query on recognition after reviewing the semantic structure of the database, and later evaluating the impact of the former on the latter was the initial concern. The essence is channeled towards modeling the information seekers subject, order than just selecting the presented keywords for document search. Thus, the model can intelligently learn and extract concept from the user's search query to further provide relevant search results for each differentiated end-users from the inferred user query concept. Beyond this, is the process of using Latent Sematic Analysis with the newly generated concept to build semantic similarities in potential retrievable documents with the view to spam documents with a higher level of cohesion and lower entropy rate. No doubt this process has outperformed the existing methods of spoken text detection and query expansion approach for relevant document retrieval in voice information retrieval system for document.

Reference:

1. Christopher D. Manning, Prabhakar Raghavance and Hinrich Schutze, Inroduction to Information Retrieval, Cambridge University Press. 2008
2. Maryam Habibi and Andrei Popescu-Beli,"Keyword extraction and clustering for Documents Recommendation in Conversation" in Proc. IEEE/ACM transaction on audio, speech, and language processing, VOL.23, NO.4 April 2015.
3. Zaslavsky, A., C. Perera and D. Georgakopoulos. Sensing as a service and Big Data. Proceedings of the International Conference on Advances in Cloud Computing, (ACC' 12), Bangalore, India, pp: 1-8, 2012.
4. Sugimoto K., Nishizaki H., and Sekiguchi Y, "Effect of document expansion using web documents for

spoken documents retrieval," in Proceedings of the 2nd Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, pp. 526–529, 2010.

5. Onifade, O.F.W., Thiery O., Osofisan, A.O. & Duffing G. "Ontological Framework for Minimizing the Risk of Non Quality Data during Knowledge Reconciliation in Economic Intelligence". In proc of the 13th Int'l Conference in Information Quality (ICIQ'08), MIT, Boston, pp 296 – 309, 2008.
6. Fayolle J., Saraclar M, Moreau F., C. Raymond, and G. Gravier, "Lexical-phonetic automata for spoken utterance indexing and retrieval," in Proceedings of International Conference on Speech Communication and Technology, 2012.
7. Paolo Rosso, Llu´ıs-F. Hurtado and Encarna Segarra,Emilio Sanchis,"On the Voice-Activated "Question Answering, IEEE Transactions on Systems, Man, and Cybernetics—part c: applications and reviews, vol. 42, no. 1,January 2012.
8. Parlak S. and Saraclar M., "Spoken term detection for Turkish broadcast news," submitted to ICASSP, 2008.
9. Wiqas Ghai and Navdeep Singh, ―Literature Review on Automatic Speech Recognition, ‖International Journal of Computer Applications (0975 – 8887) Volume 41– No.8, March 2012. [4] Y. Wang, D. Yu, Y. Ju and A. Acero, An introduction to voice search,'' IEEE Signal Process. Mag., vol. 25, no. 3, pp. 28–38, May 2008.
10. Crestani, F., Du, H.: Written Versus Spoken Queries: A Qualitative and Quantitative Comparative Analysis. J. Am. Soc. Inf. Sci., 57: 881890, 2006.
11. Hazen T, Shen W and White C "Query-by-example spoken term detection using phonetic posteriogram templates@ Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding. Merano, Italy. 2009.
12. Mamou J., Ramabhadran B., and Siohan O., "Vocabulary independent spoken term detection," in Proc. SIGIR, 2007.
13. Min Ma, Keelan Evanini, Anastassia Loukina, Xinhao Wang, Using F0 Contours to Assess Nativeness in a Sentence Repeat Task. Accepted by INTERSPEECH, Dresden, 2015.
14. Haizhou Li, Bin Ma, and Kong Aik Lee, "Spoken Language Recognition: From Fundamentals to Practice" Proceedings of the IEEE | Vol. 101, No. 5, May 2013.
15. Niladri Sekhar Dey,Ramakanta Mohanty, K.L.chugh, ‖Speech and Speaker Recognition System using Artificial Neural Networks and Hidden Markov Model‖, preceedings of International Conference on Communication Systems and Network Technologies, 2012.
16. Santosh K.Gaikwad, Bharti W.Gawali and Pravin Yannawar, "A Review on Speech Recognition Technique, International Journal of Computer Applications" Vol. 10-No.3, November 2010.
17. Gerhard Weikum, Gjergji Kas neci, Maya Ramanath, and Fab ian Suchanek, "Database and Information-Retrieval Methods for Knowledge Discovery",Communications of the ACM, vol. 52, no. 4,. April 2009
18. Chadawan Ittichaichareon, Siwat Suksri and Thaweesak Yingthawornsuk, "Speech Recognition using MFCC", International Conference on Computer Graphics, Simulation and Modeling (ICGSM'2012), Pattaya (Thailand), July 28-29, 2012.
19. Schalkwyk, J., Beeferman, D., Beaufays, F., Byrne, B., Chelba, C., Cohen, M., Strope, B: "Your Word is my Command: Google Search by Voice: A Case Study. Advances in Speech Recognition", pp. 61-90. Springer US, 2010.
20. Guozhen An, David Guy Brizan, Min Ma, Michelle Morales, Ali Raza Syed, Andrew Rosenberg, Automatic Recognition of Uni- fied Parkinsons Disease Rating from Speech with Acoustic, iVector and Phonotactic Features. Accepted by INTERSPEECH, Dresden, 2015.
21. Katsurada K., Sawada S., Teshima S., Iribe Y., and Nitta T., "Evaluation of fast spoken term detection using a suffix array," in Proceedings of International Conference on Speech Communication and Technology, pp. 909–912, 2011.
22. Itoh Y., Iwata K., Ishigame M., Tanaka K., and Shi-wook L., "Spoken term detection results using plural subword models by estimating detection performance for each query," in Proceedings of International Conference on Speech Communication and Technology, pp. 2117–2120. 2011.
23. Akiba T. and Honda K., "Effects of query expansion for spoken document passage retrieval," in Proceedings of International Conference on Speech Communication and Technology, pp. 2137–2140, 2011.
24. Tomoko T and Akiba T "Open Vocabulary Spoken Content Retrieval by front ending with spoken text detection". 2009
25. Mamou J., Ramabhadran B., and Siohan O., "Phonetic query expansion for spoken term detection" Proceedings of Interspeech. 2008
26. Toth B, Hakkani-Tur D and Yaman S Summarization and learning based approaches to information distillation. Proceedings of ICASSP, 2010.
27. Tonio Wandmacher "Hoe Semantic, is Latent Semantic Analysis?" Recital, page 6 – 10, 2005.
28. Erik Tjong Kim Sang and Katja Hofmann, "Lexical Patterns or Dependency Patterns: Which Is Better for Hypernym Extraction?" Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL) page 174- 182, 2009
29. Zheng Chen, Shengping Liu, Liu Wenyin, Geguang Pu, and Wei-Ying Ma, "Building a web thesaurus from web link structure" Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval. Pages 48 – 55, 2003
30. Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. Proceedings of the Empiricial Methods in Natural Language Processing (EMNLP 2014), 12:1532–1543, 2014.
31. Justin Richards and Min Ma and Andrew Rosenberg. Using Word Burst Analysis to Rescore Keyword Search Candidates on Low resource Languages.

Proceedings of the 39th annual international IEEE ICASSP conference. 2014.
32. Olufade F. W Onifade and Ayodeji O.J Ibitoye "Conceptual Voice Based Querying Support Model for Relevant Document Retrieval" Proceedings of the 6th International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA 2015) pp 491-497, held in Kochi, India during December 16-18, 2015.

Authors' Profiles

Olufade F.W Onifade obtained a PhD in computer science from Nancy 2 University, Nancy, France in 2009. He is currently a Senior Lecturer at the Computer Science department, University of Ibadan, Ibadan, Nigeria. He has published over 70 papers in both local and International referred journals and conferences and has held several fellowships including ETTMIT and the CV Raman Fellowship for African Researchers in India. His research interests include Fuzzy Learning, Information Retrieval, Biometrics and Pattern Matching. Dr. Onifade is a member of IEEE, IAENG and CPN.

Ayodeji O.J Ibitoye lectures at the Department of Computer Science and Information Technology, Bowen University, Iwo, Nigeria. He obtained his B.Sc. and M.sc Computer Science from the prestigious University of Ilorin, Ilorin Kwara State and University of Ibadan, Ibadan, Oyo state respectively. He is a young innovative and resourceful researcher with great analytic and programming skills. His Research Interest is in Big Data Analytics, Information Retrieval, Biometrics and Fuzzy learning. He has several peered reviewed journals and conferences in his field of expertise.