International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 10
October 2017

Caller Identification from Semantic Analysis of Call History

Kiran Mayee
Talla Padmavathi College of Engineering

ABSTRACT

Many a time we speak once or twice to one or more people about a specific topic. It is impossible to save every person's name and then remember the same in the concerned context. For example, there can be a Dev related to plumbing and another related to computer repairs. Assuming both are saved with the same name and different surnames, it becomes difficult to remember after a long break as to which one of these two is the plumber and which is the computer guy. It is in this context that speech to text conversion comes in handy. This paper deals with the solution to this problem using an inbuilt semantic knowledgebase to store keywords pertaining to previous conversations as the base to create a semantic network that would then be used to recognize and differentiate our contacts.

Keywords

Dictionary, Binary Search Tree, Speech to text, Topic Summarization

INTRODUCTION

All of us invariably use cell phones for making calls. For many such callers, we tend to store the numbers that are important to us so that we can call them back when the need arises. For users who tend to get too many calls back-to-back it becomes a hassle to remember and store the names with the intent of the caller. The same can be said of agents in intelligence, who need to look at exponential quantities of audio data pertaining to phone calls to identify potential terrorists. Their task could become easier if they were to have a ready-made semantic knowledgebase with a built-in query processor that would give them the cell numbers (and if possible, other details as well) of the people involved in coded conversations.

Many an app has been developed that show us the caller-ids whenever we dial a number or receive a call such as TrueCaller. But, these apps invade the privacy of the cell phone user, having access to all their contacts and SMSes, the same being utilized online. Also, there is an inherent danger of the app overriding all other user-friendly apps of the cell phone thereafter.

It is therefore imperative that we use an inbuilt offline app that would pick the important topics in a conversation such as the person being spoken to (assuming a formal conversation), the major topics focused on (using a summarization tool) and an option to integrate this stored knowledge into a user-friendly query processor so as to enable an optimal use of the knowledge.

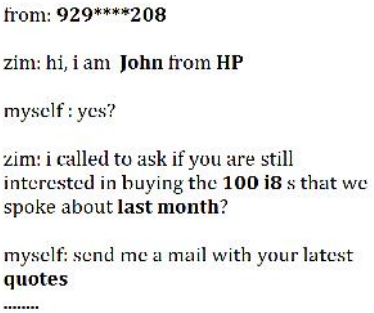

Fig 1: A typical day-to-day conversation

For example, consider the conversation presented in Figure1. The “learning table” would be as in row1 of teble1. Similarly, given the conversation of a cell phone user with zis plumber, the possible key entries in the “learning table” would be as in row two. Similarly, the key terms for the same person’s conversation with his

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 10
October 2017

boss would be as in row three. Row three is also an example where the name of the callee is missing (sir). These entries, when arranged as a semantic knowledgebase and linked to an artifact knowledgebase such as PurposeNet, would be useful to assess the name and purpose of the call, for future reference.

Table 1. A typical arrangement of the caller dictionary

Cell Number	Name	Topic	Keyword1	Keyword2s
929****208	John	Laptop	sell	100
903172XXXX	Dev	leaking pipe	repair	Saturday
937632XXXX	Sir	Report submission	negative	-----

LITERATURE SURVEY

Wagner [1] discusses communication real-time access systems (CART) with the need, methods and challenges thereof [2]. Allahyri et. al [3] discuss the various state-of-the-art techniques being used currently for summarization – topic representation approaches, frequency-driven methods, graph-based and machine learning techniques. Ayush et. al [9] use extraction based techniques for automatically generating summaries from documents using k-means clustering. Balakrishna et. al [5] discuss Polaris – a semantic parser that automatically extract deep semantic information from text based on 26 semantic relations by domain ontology creation using Jaguar with promising results. Kiran Mayee et. al[6] discuss PurposeNet – an extensive knowledgebase based on purpose as the principle of creation. Traditional approaches [8][9] to developing automatic speaker diarization tools involve supervised machine learning, which centers around having the computer program learn from annotated conversations that indicate when different speakers enter a conversation and how many are speaking at a certain time. Stephen et. al [7] have developed a probabilistic approach to speaker clustering by applying Gaussian Mix model to principal component analysis and performing iterative optimization. Susanne [10] discusses the challenges of real-time intralingual speech-to-text conversions. Bonnie et. al [11] have proposed a Universal Networking Language to overcome the inter-lingual text conversion barrier.

METHODOLOGY

Our methodology of identifying a caller cell number takes the previous conversations as input and therefore requires a copy of the call audio files. It is expected that the audio files are of reasonable clarity amenable for identifying the unique voices (speaker diarization) therein and text conversion by the existing standard audio-to-text tools. In case a real-time tool is being used for the purpose, our task becomes easier.

The process of determining the cell number of a caller from caller history follows the following major steps (Figure 2) –

(1) Creation of a semantic knowledgebase of previous calls

(2) Development of a keyword-based Query system to extract cell numbers involved in ‘topic’, prioritizing and presentation to user.

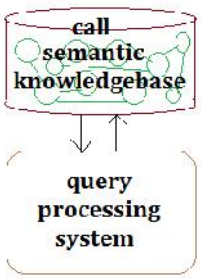

Fig 2: Major steps in Caller Identification based on call history

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 10
October 2017

Step1 is a complex process involving the determination of the various voices in a call audio file, and the extraction of key words in the conversation. These keywords are used to create a semantic knowledgebase with the caller number intact. Step2 involves the development of a user-friendly query processor that takes any keyword from the user, searches the knowledgebase for calls that were made with that keyword(s), uses prioritization to determine which among the cell numbers extracted is to be presented to the user, thereby allowing the user to recall anyone who has been in a call in history without having to remember the name(s).

SEMANTIC KNOWLEDGEBASE CREATION

The caller-identification tool involves six tasks – conversation saving and retrieval module (in case of offline processing), speaker diarization (recognition and differentiation), speech-to-text conversion, keyword extraction, text summarization, semantic network development and linking to PurposeNet. (Figure 1).

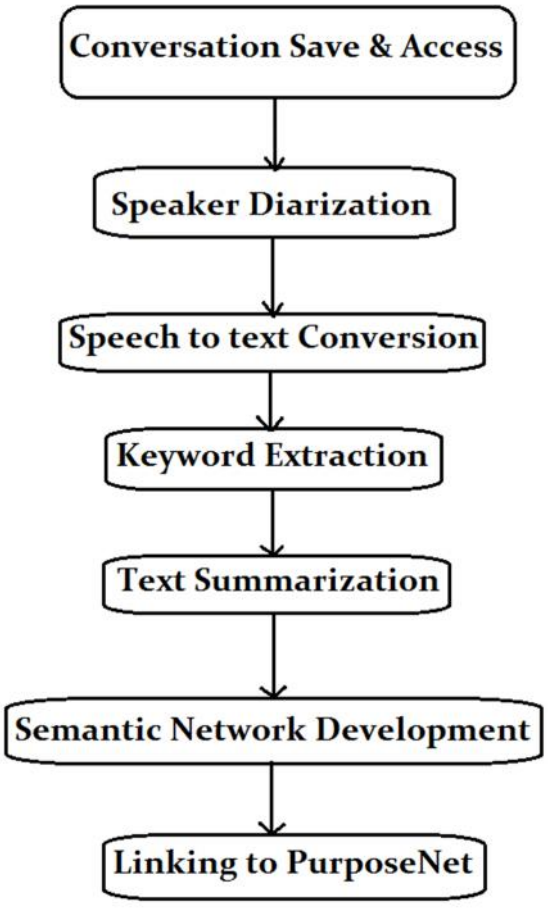

Fig 3: Methodology to convert a Call to Semantic Knowledgebase

A PROPOSAL FOR THE KNOWLEDGEBASE DESIGN

We propose the following knowledgebase that follows closely on lines of [12][13] with the basic entries – Cell Number , Date-time, MyKeywords, ZisKeywords. Keywords could be from the Polaris semantic set of 26 and more. For example, purpose-topic, artifact, action, time-of-action, etc. The basic set of attributes that are extracted and inserted in the knowledgebase for every call made are - the cell number, name (can be missing), affiliation (from) and purpose. Figure 4 shows the knowledgebase created from keywords extracted from the conversation given in figure 1.

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 10
October 2017

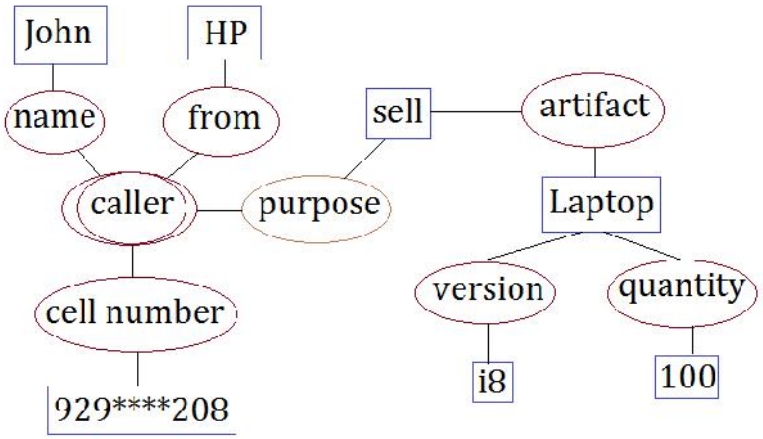

Fig 4: Semantic knowledgebase created from conversation in figure 1

ISSUES

Differentiating the two or more voices in a call is a difficult process even with the current technology. Many modules are required for the production of a semantic knowledgebase carrying the important information in a call. In case of confidential calls, the information in the 'dictionary' too is confidential and therefore, there need to be a module to allow the user to "lock" such conversations and "hide" such dictionaries. Most conversations today are inter-lingual. Therefore, the speech to text converter has to have a module accordingly. Finally, there is also a possibility of the proposed tool and semantic dictionary needing considerable space, which would definitely be a deterrent for today's cell phones.

CONCLUSION

This paper tries to solve the problem of having to remember the time of every telephonic conversation or to save each number to re-communicate to the caller regarding the same topic by creating a semantic knowledgebase on the lines of PurposeNet, with the keywords spoken during the call. The same can be implemented as an app that automatically performs the background processing either in real-time or as an offline process using the call audio file as input. The same can be accessed using a simple query system by the user of the cell phone, thereby exponentially reducing the human memory load and time as well. That said, research in this area is still at its nascent stage and has a lot of scope for further research.

REFERENCES

[1] Khilari, Prachi, Bhop.V.P. 2015. A Review on Speech to text Conversion Methods. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 7, July 2015.
[2] Wagner, Sussanne. 2005. Intralingual Speech-to-text conversion in real-time: Challenges and Opportunities. MuTra 2005 – Challenges of Multidimensional Translation: Conference Proceedings.
[3] Allahyari, Mehdi, Seyedamin Pouriyeh, Assefi Mehdi, Saeid Safaei, Elizabeth D. Trippe, Gutierrez Juan B., Kochut Krys. Text Summarization Techniques: A Brief Survey. arXiv: 1707.02268v3 [cs.CL] 28 Jul 2017.
[4] Real-Time Speech Emotion and Sentiment Recognition for Interactive Dialogue System. Proceedings of the 2016 ACM Conference on Empirical Methods in Natural Language Processing. Pp. 1042-1047 November 1-5, 2016.
[5] Balakrishna Mithun, Moldovan Dan, Tatu Marta, Olteanu Marian. Semi-automatic domain ontology creation from text resources. Pp. 3187- 3194.
[6] Mayee Kiran. PurposeNet. Lambert academic publishing. 2016.
[7] S.Tranter and D. Reynolds. An overview of automatic speaker diarisation systems. IEEE Trans. Audio, Speech, Lang. Process. vol.14, no. 5, pp. 1557–1565, Sep. 2006.

[7] M. Moattar and M. M. Homayounpour. A review on speaker diarization systems and approaches. Speech Commun. vol. 54, no. 10. pp. 1065–1103, 2012.
[8] X. Anguera, S. Bozonnet, N. Evans, C. Fredouille, G. Friedland, and O. Vinyals. Speaker diarization: A review of recent research. IEEE Trans. Audio, Speech, Lang. Process. vol. 20, no. 2, pp. 356–370. 2012.
[9] Agrwal Ayush, Gupta Utsav. Extraction v=based approach for text summarization using k-means clustering. International Journal of scientific and research publications. Vol. 4, Issue 11. 2014.
[10] Wagner, Susanne. Intralingual speech-to-text conversion in real-time: Challenges and opportunities. Conference Proceedings of MuTra 2005 – Challenges of Multidimensional Translation. 2004.
[11] Dorr Bonnie J., Hovy Eduard H., Levin Loris S. Machine Translation : Interlingual Methods. In Natural Language Processing and Machine Translation Encyclopedia of Language and Linguistics, 2nd ed. (ELL2).
[12] Bateman, J.A., Kasper, R.T., Moore, J. D., and Whitney, R.A. A General organization of knowledge for Natural Language Processing. The Penman upper model. Research Report. USC/Information Sciences Institute. CA. 1989.
[13] Nirenburg, S. and V.Raskin. Ontological Semantics. Cambridge: MIT press. 2004.
[14] Mayee Kiran. P, Sangal Rajeev, Paul Soma. Action semantics in PurposeNet. Proceedings of IEEE World Congress on Information and Communication Tecnologies. 2011.